

Bayesian inference based on algorithms: MH, HMC, MALA and LIP-MALA for Prestack Seismic Inversion

Richard Perez-Roa¹, Saba Infante^{2,3}, Gabriel Barragan², Raul Manzanilla²

¹Geology and Applied Geophysics Research Group, School of Earth Sciences, Energy and Environment, Yachay Tech University, Hacienda San José, Urcuquí, 100119, Ecuador

5 ²Numerical Analysis and Data Science Research Group, School of Mathematical and Computational Sciences, Hacienda San José, Urcuquí, 100119, Ecuador

³Department of Mathematics, FACyT Universidad de Carabobo, Naguanagua, Valencia, 2005, Venezuela

Correspondence to: Richard Perez-Roa (rperez@yachaytech.edu.ec)

10 **Abstract.** Seismic inversion for estimating elastic properties is a key technique for reservoir characterization after drilling. The choice of inversion method strongly influences the accuracy, efficiency, and reliability of results. Bayesian inference based on Markov Chain Monte Carlo (MCMC) algorithms provides a robust framework for incorporating data uncertainty and prior geological knowledge. In this study, we compare the performance of four inversion methods—Metropolis-Hastings (MH), Hamiltonian Monte Carlo (HMC), the Metropolis-Adjusted Langevin Algorithm (MALA), and its variant Lip-
15 MALA—in prestack seismic inversion using both synthetic models and real data from an eastern Venezuelan hydrocarbon reservoir. Results indicate that gradient-based methods (HMC, MALA, Lip-MALA) outperform MH in velocity estimation, while density inversion remains more challenging. MH and MALA achieve shorter execution times, whereas HMC and Lip-MALA improve accuracy at higher computational cost. This analysis evaluates mean values and standard deviation (SD) estimates for P-wave velocity, S-wave velocity, and density, with quality assessed through correlation metrics, objective func-
20 tion behavior, seismic traces, and Root Mean Square Error (RMSE). A two-dimensional inversion with real data further demonstrates algorithms performance under complex geological conditions. The findings highlight trade-offs between accuracy and efficiency, providing practical guidelines for selecting inversion method in seismic reservoir characterization.

1 Introduction

The accurate characterization of hydrocarbon reservoirs is fundamental for effective exploration and production decisions.
25 This requires integrating complementary sources of information: general geological knowledge from analogous reservoirs and rock physics principles, and reservoir-specific observations from well logs, seismic surveys, and production history.

Seismic data play a central role due to their extensive spatial coverage, providing a continuous description of the subsurface beyond individual well locations. To translate seismic amplitudes into meaningful elastic and petrophysical properties, such as P-wave velocity (V_p), S-wave velocity (V_s), and density (ρ), seismic inversion is employed. This geophysical inverse prob-
30 lem aims to infer subsurface properties from observed seismic data, typically expressed as (Tarantola, 2005):

$$\mathbf{d}_{obs} = g(\mathbf{m}) + \xi, \quad (1)$$

where \mathbf{d}_{obs} are the observed data, g is the forward operator linking the model parameters \mathbf{m} to the data, and ξ represents measurement and modeling errors.

35 Traditional seismic inversion methods can be broadly categorized into deterministic and stochastic approaches. Deterministic methods, such as least-squares inversion (Tarantola, 2005) and gradient-based optimization (Buland and Omre, 2003), provide single best-fit models but often lack robust uncertainty quantification. Stochastic methods, including simulated annealing (Ma, 2002) and genetic algorithms, explore the model space more broadly but can be computationally intensive and may not guarantee convergence to the posterior distribution.

40 In prestack amplitude versus offset (AVO) inversion (Helland-Hansen et al., 1997; Ma, 2002; Buland and Omre, 2003), the problem is ill-posed (Landa and Treitel, 2016): solutions are non-unique and highly sensitive to data noise and modeling errors (Landa and Treitel, 2016). The elastic parameters (V_p , V_s , ρ), from which Lamé parameters can be derived. These are sensitive to rock fluid content and saturation (Clochard et al., 2009) and can be further related to petrophysical parameters such as porosity, sand/shale ratio, and gas saturation (Goodway, 2001). Accurately estimating these parameters is therefore crucial for identifying hydrocarbon accumulations.

45 To address these challenges, Bayesian inference has emerged as a powerful framework, as it explicitly incorporates prior geological knowledge and quantifies uncertainty through the posterior distribution. Markov Chain Monte Carlo (MCMC) methods are particularly well-suited for sampling complex, high-dimensional posterior distributions in nonlinear inverse problems (Bosch et al., 2007).

50 Several MCMC algorithms have been applied in geophysical inversion. The Metropolis–Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) is a foundational method but can suffer from slow convergence in high-dimensional spaces. Hamiltonian Monte Carlo (HMC) (Duane et al., 1987; Neal, 2012) leverages gradient information to propose more efficient moves, making it suitable for high-dimensional problems, as demonstrated in full-waveform inversion (Gebraad et al., 2020) and acoustic inversion (de Lima et al., 2023). Langevin dynamics-based methods, such as the Metropolis-Adjusted Langevin Algorithm (MALA) (Roberts and Tweedie, 1996) and its adaptive variant Lip-MALA (Nemeth and Fearnhead, 2021; Izzatullah et al., 2021), use gradient information to guide proposals, improving mixing and convergence in complex posteriors.

55 Despite these advances, a systematic comparison of these MCMC algorithms in the context of prestack seismic inversion—particularly using both synthetic and real data in 1D and 2D settings—remains limited. Most studies focus on a single algorithm or synthetic cases, leaving a gap in understanding their practical performance under realistic geological complexity and computational constraints.

60 This study aims to fill this gap by conducting a comprehensive evaluation of four MCMC algorithms—MH, HMC, MALA, and Lip-MALA—for Bayesian prestack seismic inversion. We assess their performance in estimating elastic parameters (V_p , V_s , ρ) using both noise-free synthetic data and real data from an eastern Venezuelan reservoir. Our analysis includes:

- Quantitative metrics: mean, standard deviation, correlation, and RMSE.
- Computational efficiency: acceptance rates and execution times.
- 65 • Convergence diagnostics: multivariate effective sample size (mESS).
- Practical applicability: extension to 2D inversion with real data.

Our contributions are threefold:

1. We provide a benchmark comparison of MCMC algorithms in seismic inversion, highlighting trade-offs between accuracy, uncertainty quantification, and computational cost.
- 70 2. We demonstrate the scalability of these methods from 1D to 2D real-data scenarios, validating their robustness in geologically complex settings.
3. We offer practical guidelines for algorithm selection based on project requirements, data characteristics, and available computational resources.

In Prestack AVO, Bayesian approaches with MCMC have been previously applied to estimate impedance or elastic parameters (e.g., Bosch et al., 2007; Wu et al., 2019), but these studies focus on a single algorithm or mainly on synthetic cases, but we go further and test it on real data.

The remainder of this paper is organized as follows: Section 2 formulates the seismic inversion problem within a Bayesian framework and describes the forward modeling approach based on AVO theory. Section 3 details the theoretical background of the MCMC algorithms. Section 4 presents result from synthetic and real data experiments. Section 5 discusses the implications and practical recommendations. Section 6 concludes the study.

2 The Seismic Inversion Problem

Having established the importance of Bayesian MCMC methods for prestack seismic inversion in Section 1, this section provides a comprehensive formulation of the inverse problem within a rigorous statistical framework. We detail the mathematical foundations, model parameterization, forward modeling approach, and inference workflow that form the basis of our comparative study. The complete Bayesian inference flow implemented in this study is visually summarized in Figure 1, which illustrates the steps from input data to subsequent uncertainty quantification.

2.1 Bayesian Formulation of the Inverse Problem

Within the Bayesian paradigm, the solution to the inverse problem is expressed probabilistically through the posterior distribution (Tarantola, 2005; Bosch et al., 2007) $p(\mathbf{m}|\mathbf{d}_{obs})$, which combines prior knowledge with information from the observed data. Applying Bayes' theorem, we have:

$$p(\mathbf{m}|\mathbf{d}_{obs}) = \frac{p(\mathbf{d}_{obs}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d}_{obs})} \propto cL(\mathbf{m})\rho(\mathbf{m}), \quad (2)$$

where:

- $\rho(\mathbf{m})$ is the prior distribution, representing geological knowledge independent of the seismic data.
- $L(\mathbf{m})$ is the likelihood function, quantifying the probability of observing the data given a specific model.
- c is a normalization constant.
- And $p(\mathbf{d}_{obs}) = \int p(\mathbf{d}_{obs}|\mathbf{m})p(\mathbf{m}) d\mathbf{m}$

2.1.1 Prior Distribution Specification

The prior distribution encodes constraints on the model parameters based on independent information such as well logs, regional geology, and rock physics relationships. We employ a Gaussian prior distribution $\rho(\mathbf{m}) \sim \mathcal{N}(\mathbf{m}_{prior}, \mathbf{C}_m)$ (Bosch, 2004; Buland and Omre, 2003):

$$\rho(\mathbf{m}) = c_1 \exp\left[-\frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{prior})\right], \quad (3)$$

where \mathcal{N} denotes a Gaussian probability distribution, with mean $\mathbf{m}_{prior} \in \mathbb{R}^{N_m}$, typically derived from low-frequency interpolation of well data or seismic velocities, covariance matrix given by $\mathbf{C}_m \in \mathbb{R}^{N_m \times N_m}$, which imposes spatial correlation and smoothness constraints, effectively regularizing the ill-posed inverse problem and c_1 an appropriate normalization constant.

2.1.2 Likelihood Function and Data Errors

Assuming Gaussian-distributed errors in Eq. (1) with zero mean and covariance matrix \mathbf{C}_d , the likelihood function takes the form:

$$L(\mathbf{m}) = \exp\left[-\frac{1}{2}(\mathbf{d}_{obs} - g(\mathbf{m}))^T \mathbf{C}_d^{-1}(\mathbf{d}_{obs} - g(\mathbf{m}))\right], \quad (4)$$

The likelihood function assumes Gaussian-distributed errors, consistent with standard formulations in Bayesian seismic inversion (Tarantola, 2005; Izzatullah et al., 2021), $g: \mathbf{m} \rightarrow \mathbf{d}_{obs}$, being the function solving the seismic forward problem and the

data covariance matrix \mathbf{C}_d accounts for measurement noise and uncertainties in the forward model. While often assumed diagonal (uncorrelated errors), it can incorporate more complex error structures when available.

2.1.3 Posterior Distribution and MCMC Sampling

Combining Eqs. (3) and (4), the posterior distribution can be expressed through the negative log-posterior (misfit function) $S(\mathbf{m})$:

$$S(\mathbf{m}) = -\log p(\mathbf{m}|\mathbf{d}_{obs}) = \frac{1}{2}(\mathbf{d}_{obs} - g(\mathbf{m}))^T \mathbf{C}_d^{-1}(\mathbf{d}_{obs} - g(\mathbf{m})) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{prior}), \quad (5)$$

Thus,

$$p(\mathbf{m}|\mathbf{d}_{obs}) \propto \exp[-S(\mathbf{m})], \quad (6)$$

In practice, the posterior distribution defined above is analytically intractable. Markov Chain Monte Carlo (MCMC) methods are therefore employed to generate samples $\{\mathbf{m}^{(i)}\}_{i=1}^N$ from Eq. (6), enabling comprehensive uncertainty quantification through estimation of posterior means, variances, and marginal distributions (Robert and Casella, 2004; Neal, 2012; Gebraad et al., 2020).

2.2 Model Parameterization and Forward Modelling

2.2.1 Logarithmic Parameterization

To ensure positivity of physical parameters ($V_p > 0$, $V_s > 0$, $\rho > 0$) and improve numerical stability, we adopt a logarithmic parameterization. For a 1D model with N layers, the parameter vector is (Buland and Omre, 2003):

$$\mathbf{m} = [\ln(V_{p,1}), \dots, \ln(V_{p,N}), \ln(V_{s,1}), \dots, \ln(V_{s,N}), \ln(\rho_1), \dots, \ln(\rho_N)]^T, \quad (7)$$

This transformation also aligns naturally with the logarithmic form of the Aki-Richards reflectivity approximation used in our forward model.

2.2.1 AVO Forward Modeling Formulation

The forward operator $g(\mathbf{m})$ implements the Aki-Richards linearized approximation for PP reflectivity. The AVO method was created in the early 1980s to analyze the amplitudes of seismic CMP gathers as a function of angle to find hydrocarbons. The Aki-Richards equation (Aki and Richards, 2002) is the foundation of AVO analysis. The original form of the equation can be rewritten for a weak-contrast interface to give (Buland and Omre, 2003; Niu et al., 2020):

$$R_{pp}(\theta) = c_1(\theta) \frac{\Delta V_p}{V_p} + c_2(\theta) \frac{\Delta V_s}{V_s} + c_3(\theta) \frac{\Delta \rho}{\bar{\rho}}, \quad (8)$$

where

$$c_1(\theta) = \frac{1}{2}(1 + \tan^2 \theta), \quad (9)$$

$$c_2(\theta) = -4 \frac{\overline{V_s}}{\overline{V_p}} \sin^2 \theta, \quad (10)$$

140

$$c_3(\theta) = \frac{1}{2} \left(1 - 4 \frac{\overline{V_s}}{\overline{V_p}} \sin^2 \theta \right), \quad (11)$$

In equations (8 - 11), the incident angle θ is the angle at which a wave reflected off a surface. V_p, V_s and ρ represent the velocities of P-waves, S-waves, and the density of a material, respectively. $\Delta V_p, \Delta V_s$ and $\Delta \rho$ are the changes in V_p, V_s and ρ across a reflective interface. $\overline{V_p}, \overline{V_s}$ and $\overline{\rho}$ are the average values of V_p, V_s and ρ , respectively.

To obtain the seismic trace for a certain theta angle we can use the approximation for small reflectivity (Russell et al., 2006),

145

$$T(\theta) = \frac{1}{2} c_1 W(\theta) D L_{V_p} + \frac{1}{2} c_2 W(\theta) D L_{V_s} + \frac{1}{2} c_3 W(\theta) D L_{\rho}, \quad (12)$$

where $L_{V_p} = \ln(V_p)$, $L_{V_s} = \ln(V_s)$, $L_{\rho} = \ln(\rho)$, W is the wavelet matrix and D is the derivative matrix. Equation (12) can be implemented in matrix form as

$$\begin{bmatrix} T(\theta_1) \\ T(\theta_2) \\ \vdots \\ T(\theta_n) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} c_1 W(\theta_1) D & c_2 W(\theta_1) D & c_3 W(\theta_1) D \\ c_1 W(\theta_2) D & c_2 W(\theta_2) D & c_3 W(\theta_2) D \\ \vdots & \vdots & \vdots \\ c_1 W(\theta_n) D & c_2 W(\theta_n) D & c_3 W(\theta_n) D \end{bmatrix} \begin{bmatrix} L_{V_p} \\ L_{V_s} \\ L_{\rho} \end{bmatrix} \quad (13)$$

150

In discrete matrix form, for multiple angles $\{\theta_1, \theta_2, \dots, \theta_n\}$, the complete forward modeling operation can be expressed compactly as:

$$\mathbf{d}_{syn} = g(\mathbf{m}) = \mathbf{Gm}, \quad (14)$$

where:

$$\mathbf{d}_{syn} = \begin{bmatrix} T(\theta_1) \\ T(\theta_2) \\ \vdots \\ T(\theta_n) \end{bmatrix}; \quad \mathbf{G} = \frac{1}{2} \begin{bmatrix} c_1 W(\theta_1) D & c_2 W(\theta_1) D & c_3 W(\theta_1) D \\ c_1 W(\theta_2) D & c_2 W(\theta_2) D & c_3 W(\theta_2) D \\ \vdots & \vdots & \vdots \\ c_1 W(\theta_n) D & c_2 W(\theta_n) D & c_3 W(\theta_n) D \end{bmatrix} \quad (15)$$

155

and \mathbf{m} was defined in equation (7). For how to construct W and D , known as the wavelet matrix and the derivative matrix respectively, the reader is advised to review Hampson et al., (2005).

2.3 Inference Workflow and Uncertainty Quantification

The complete Bayesian inversion workflow implemented in this study is visually summarized in Figure 1. The process integrates prior geological knowledge, seismic data, physical modelling, and statistical sampling to produce probabilistic estimates of subsurface elastic properties.

160

Bayesian inversion process implemented in this study follows a systematic workflow:

1. Prior Model Construction: Define \mathbf{m}_{prior} and \mathbf{C}_m using available well-log data, geological interpretation, and spatial correlation models.
2. Data Preparation and Uncertainty Quantification: Extract angle-stacked seismic traces \mathbf{d}_{obs} , estimate corresponding wavelets for multiple angles $\{\theta_1, \theta_2, \dots, \theta_n\}$, and characterize data uncertainties through the covariance matrix \mathbf{C}_d .
- 165 3. Forward Operator Assembly: Construct the matrix \mathbf{G} according to Eqs. (8)-(15) to implement the forward modelling operator $g(\mathbf{m})$.
4. MCMC Sampling: Apply one of the four sampling algorithms (MH, HMC, MALA, or Lip-MALA) to generate a Markov chain of model samples from the posterior distribution $p(\mathbf{m}|\mathbf{d}_{obs})$.
- 170 5. Posterior Analysis: Compute statistical summaries (posterior means, standard deviations, correlation coefficients, RMSE) and marginal probability distributions from the sampled models to assess the inverted elastic parameters and their associated uncertainties.

This comprehensive framework provides a rigorous approach to seismic inversion where solutions are characterized not as single models but as probability distributions that fully quantify uncertainty, directly addressing the ill-posed nature of the inverse problem.

175 **3 Theoretical Background for Metropolis-Hastings, Hamiltonian Monte Carlo and Langevin Diffusion**

This section presents the theoretical foundations of the four MCMC algorithms compared in this study. A visual summary of their computational workflows and key characteristics is provided in Figure 2, followed by detailed mathematical descriptions and implementation considerations.

3.1 Algorithm Overview and Comparative Framework

- 180 Figure 2 provides a schematic comparative view of the four MCMC algorithms, highlighting their proposal mechanisms, gradient requirements, and tuning strategies. All methods share the common goal of generating samples from the posterior distribution $p(\mathbf{m}|\mathbf{d}_{obs})$ but employ different strategies for proposing and accepting new states.

3.2 Metropolis–Hastings Algorithm

- 185 MH algorithm is one of the most widely used MCMC methods for sampling from complex probability distributions. Originally proposed by Metropolis et al. (1953) and generalized by Hastings (1970), MH defines a transition probability that ensures ergodicity, detailed balance, and reversibility of the chain (Chib and Greenberg, 1995).

3.2.1 Mathematical Formulation

Given the current state $\mathbf{m}^{(t)}$ at iteration t , MH proposes a new candidate \mathbf{m}^* from a proposal distribution $q(\mathbf{m}^*|\mathbf{m}^{(t)})$. The candidate is accepted with probability:

$$\alpha_{MH} = \min \left\{ 1, \frac{L(\mathbf{m}^*)}{L(\mathbf{m}^{(t)})} \cdot \frac{q(\mathbf{m}^{(t)}|\mathbf{m}^*)}{q(\mathbf{m}^*|\mathbf{m}^{(t)})} \right\}, \quad (16)$$

where $L(\cdot)$ represents the likelihood function (Eq. 4). For symmetric proposal distributions where $q(\mathbf{m}^*|\mathbf{m}^{(t)}) = q(\mathbf{m}^{(t)}|\mathbf{m}^*)$, this simplifies to the Metropolis acceptance ratio.

3.2.2 Implementation Details

In our seismic inversion context, we employ a Gaussian random walk proposal:

$$q(\mathbf{m}^*|\mathbf{m}^{(t)}) = \mathcal{N}(\mathbf{m}^*|\mathbf{m}^{(t)}, \sigma^2\mathbf{I}), \quad (17)$$

Where σ is a tuning parameter controlling the proposal step size. The algorithm proceeds as follows:

- Step 1. Initialization: Start with an initial model $\mathbf{m}^{(0)} = \mathbf{m}_{prior}$ and set $t = 0$.
- Step 2. Propose a new state: Generate $\mathbf{m}^* \sim \mathcal{N}(\mathbf{m}^{(t)}, \sigma^2\mathbf{I})$.
- Step 3. Compute Acceptance: Calculate α_{MH} using Eq. (16).
- 200 Step 4. Accept/Reject: Draw $u \sim Uniform(0,1)$. If $u \leq \alpha_{MH}$, set $\mathbf{m}^{(t+1)} = \mathbf{m}^*$; otherwise $\mathbf{m}^{(t+1)} = \mathbf{m}^{(t)}$.
- Step 5. Iterate: Increment t and repeat from step 2 until convergence.

3.2.3 Tuning Considerations and Limitations

The performance of MH critically depends on the choice of σ . Optimal scaling theory suggests aiming for acceptance rates of approximately 23.4% for random walk proposals in high dimensions (Roberts et al., 1997). In our implementation, σ was
205 tuned during a 1000-iteration burn-in period to achieve this target.

While MH is conceptually simple and requires no gradient computations, its random walk behavior leads to slow exploration of high-dimensional parameter spaces. The autocorrelation between samples remains high, requiring longer chains to achieve effective sample sizes comparable to gradient-based methods.

3.3 Hamiltonian Monte Carlo (HMC)

210 HMC is a sampling algorithm that was first introduced in molecular dynamics (Duane et al., 1987) and later adapted to Bayesian inference problems (Neal, 2012, Fichtner and Zunino, 2019). It is particularly effective for high-dimensional

posteriors when gradient information of the σ is available and it is more efficient than standard MH. It introduces an auxiliary momentum variable to simulate Hamiltonian dynamics in an augmented parameter space.

3.3.1 Hamiltonian Dynamics Formulation

215 HMC is an MCMC algorithm that uses classical Hamiltonian mechanics (Landau and Lifshitz, 1976) to sample from an arbitrary n -dimensional probability density function (PDF). HMC introduces momentum variables $\mathbf{p} \in \mathbb{R}^{N_m}$ and samples from the joint distribution:

$$p(\mathbf{m}, \mathbf{p}) = p(\mathbf{m})p(\mathbf{p}) \propto \exp(-U(\mathbf{m}) - K(\mathbf{p})) \quad (18)$$

where $U(\mathbf{m}) = -\log(p(\mathbf{m}))$, which is defined as the negative logarithm of the PDF (Gebraad et al., 2020), and

220 $K(\mathbf{p}) = \frac{1}{2}\mathbf{p}^T \mathbf{M}^{-1}\mathbf{p}$ is the kinetic energy with mass matrix \mathbf{M} .

The Hamiltonian equations of motion are:

$$\frac{d\mathbf{m}}{d\tau} = \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p}, \quad \frac{d\mathbf{p}}{d\tau} = -\frac{\partial H}{\partial \mathbf{m}} = -\nabla U(\mathbf{m}) \quad (19)$$

where $H(\mathbf{m}, \mathbf{p}) = U(\mathbf{m}) + K(\mathbf{p})$ is the Hamiltonian and τ is fictitious time.

3.3.2 Leapfrog Integration and Algorithm

225 Since analytical solutions to Eq. (19) are unavailable for nonlinear problems, HMC employs the symplectic leapfrog integrator with step size ϵ and L steps. For our seismic inversion, the algorithm proceeds as (Gebraad et al., 2020):

Step 1. Start with an initial model $\mathbf{m}^{(0)} = \mathbf{m}_{prior}$, random momenta $\mathbf{p}^{(0)} \sim \mathcal{N}(0, \mathbf{M})$. and set $t = 0$.

Step 2. Evaluate the Hamiltonian $H(\mathbf{m}^{(t)}, \mathbf{p}^{(t)})$ starting with $t = 0$.

Step 3. Leapfrog integration (repeated L times): Calculate:

230
$$\mathbf{p}^* = \mathbf{p}^{(t)} - \frac{\epsilon}{2}\nabla U(\mathbf{m}^{(t)}) \quad (20)$$

$$\mathbf{m}^* = \mathbf{m}^{(t)} + \epsilon\mathbf{p}^* \quad (21)$$

$$\mathbf{p}^* = \mathbf{p}^* - \frac{\epsilon}{2}\nabla U(\mathbf{m}^*) \quad (22)$$

Step 4. Calculate the Hamiltonian $H(\mathbf{m}^*, \mathbf{p}^*)$.

Step 5. Compute Acceptance: Calculate α_{HMC} using:

235
$$\alpha_{HMC} = \min\left(1, \exp(H(\mathbf{m}^{(t)}, \mathbf{p}^{(t)}) - H(\mathbf{m}^*, \mathbf{p}^*))\right), \quad (23)$$

Step 6. Accept/Reject: Draw $u \sim \text{Uniform}(0,1)$. If $u \leq \alpha_{HMC}$, set $\mathbf{m}^{(t+1)} = \mathbf{m}^*$ and $H(\mathbf{m}^{(t+1)}, \mathbf{p}^{(t+1)}) = H(\mathbf{m}^*, \mathbf{p}^*)$; otherwise $\mathbf{m}^{(t+1)} = \mathbf{m}^{(t)}$ and $\mathbf{p}^{(t+1)} \sim \mathcal{N}(0, \mathbf{M})$.

Step 7. Iterate: Increment t and repeat from step 2 until convergence.

The acceptance rate of the leapfrog integration algorithm commonly used in Step 3 is largely influenced by how well it con-
 240 serves energy in the trajectory. If the time steps are too large or the gradients of the fitting function are incorrectly calculated,
 the algorithm will save less energy, and the acceptance rate will decrease. Simply put, the leapfrog integration algorithm works
 by bouncing model parameters back and forth across the simulated energy landscape. The acceptance rate determines how
 often the algorithm accepts a new proposed model parameter. If the time steps are too large or the gradients are calculated
 incorrectly, the algorithm cannot follow the energy landscape accurately and will likely reject the proposed model parameters.
 245 This leads to slower convergence and increases computational cost.

3.3.3 Parameter Tuning for Seismic Inversion

HMC requires careful tuning of two parameters: the step size ϵ and trajectory length L . We employed a dual adaptation strat-
 egy:

1. Step size (ϵ): Tuned during burn-in to achieve acceptance rates optimal for HMC.
- 250 2. Trajectory length (L): Fixed at $L = 10$ based on preliminary experiments showing adequate exploration without ex-
 cessive computational cost.

3.4 Langevin Diffusion-Based Algorithms

3.4.1 Langevin Dynamics

Langevin dynamics are a mathematical model of Brownian motion, named after the French physicist Paul Langevin (Lemons
 255 and Gythiel, 1997) who developed them in 1908. Langevin dynamics is a simplification of Albert Einstein's approach to
 Brownian motion, which is based on Newton's second law of motion. The Langevin dynamics for target distribution $\sigma(\mathbf{m}_t)$,
 is a continuous-time stochastic process $(\mathbf{m}_t)_{t \geq 0}$ in \mathbb{R}^{N^m} that evolves following the stochastic differential equation (SDE)
 (Roberts and Stramer, 2002; Nemeth et al., 2016; Izzatullah et al., 2021) and (Sanchez et al., 2016; Infante et al., 2019):

$$d\mathbf{m}(\tau) = -\Sigma \nabla \log[p(\mathbf{m}(\tau))] d\tau + \sqrt{2\Sigma} \frac{-1}{2} d\mathbf{W}(\tau) \quad (24)$$

260 where $\mathbf{W}(\tau)$ is a standard n -dimensional Brownian motion, Σ is a symmetric positive definite matrix (in this paper we use $\Sigma =$
 \mathbf{I}), $\nabla \log[\sigma(\mathbf{m}_t)]$ is the drift term of the Brownian particle \mathbf{m}_t and $\sigma(\mathbf{m}_t)$ is a stationary posterior distribution.

3.4.2 Metropolis-adjusted Langevin algorithm (MALA)

In the practice, a standard approach is to discretize the equation (13) using the Euler-Maruyama discretization (Stuart et al.,
 2004) and we obtained the Unadjusted Langevin algorithm (ULA) given by

$$265 \quad \mathbf{m}^* = \mathbf{m}^{(t)} + \tau \Sigma \nabla \log[p(\mathbf{m}^{(t)})] + \sqrt{2\tau \Sigma} \frac{-1}{2} \epsilon_t, \quad \epsilon_t \sim N(0, \mathbf{I}_{n \times n}) \quad (25)$$

where τ is the step-length for each iteration. ULA is simple in its implementation, yet it introduces a bias, then we need to
 introduce the acceptance-rejection step through the MH algorithm. By introducing MH algorithm into ULA, we will obtain

the Metropolis-Adjusted Langevin algorithm (MALA), (Izzatullah et al., 2020, Izzatullah et al., 2021). The procedure consists of constructing a Markov chain at each step t , given $\mathbf{m}^{(t)}$, a new observation is generated from a proposal density $q(\mathbf{m}^*)$. The candidate is then accepted with probability α_{MALA} given by,

$$\alpha_{MALA} = \min \left\{ 1, \frac{L(\mathbf{m}^*)}{L(\mathbf{m}^{(t)})} \cdot \frac{q(\mathbf{m}^{(t)}|\mathbf{m}^*)}{q(\mathbf{m}^*|\mathbf{m}^{(t)})} \right\}, \quad (26)$$

In summary, MALA algorithm is obtained as follows:

- Step 1. Choose an initial solution $\mathbf{m}^{(0)} = \mathbf{m}_{prior}$ and the discretization step-length τ .
- Step 2. Draw $\epsilon_t \sim N(0, I_{n \times n})$ and simulate a new sample from the Langevin diffusion (Eq. (25)).
- Step 3. Compute the accept-reject probability α_{MALA} using Eq. (26).
- Step 4. Draw u from a uniform distribution $u \sim Uniform(0,1)$, if $\alpha_{MALA} < u$ set $\mathbf{m}^{(t+1)} = \mathbf{m}^*$; otherwise $\mathbf{m}^{(t+1)} = \mathbf{m}^{(t)}$.
- Step 5. Then, repeat this process until the convergence.

3.4.3 MALA with locally Lipschitz adaptive step size (Lip-MALA)

In the MALA algorithm, it is required to calibrate the step-size τ , because τ must decrease with dimension, n . then τ can be turned such that the MCMC achieve better mixing performance. An extension of ULA and similar in spirit with Stochastic Gradient Langevin Dynamics algorithm proposed by Welling and Teh, (2011) by suppressing the MH acceptance steps. See in (Izzatullah et al., 2021) proposed ULA variant, Lip-MALA, in which the step-length τ is adapted according to the Lipschitz condition,

$$\tau_t = \frac{1}{2} \frac{\|\mathbf{m}^{(t)} - \mathbf{m}^{(t-1)}\|_2}{\|\nabla \log[p(\mathbf{m}^{(t)})] - \nabla \log[p(\mathbf{m}^{(t-1)})]\|_2} \quad (27)$$

The general steps for Lip-MALA MCMC with locally Lipschitz adaptive step size are:

- Step 1. Choose an initial solution $\mathbf{m}^{(0)} = \mathbf{m}_{prior}$, the discretization step-length τ , $\beta_0 = +\infty$ and $L_c = N_m^{-1/3}$.
- Step 2. Draw $\epsilon_t \sim N(0, I_{n \times n})$ and simulate a new sample from the Langevin diffusion (Eq. (25)).
- Step 3. Compute the accept-reject probability α_{MALA} using Eq. (26).
- Step 4. Draw u from a uniform distribution $u \sim Uniform(0,1)$, if $\alpha_{MALA} < u$ set $\mathbf{m}^{(t+1)} = \mathbf{m}^*$, then update,

$$\tau_t = \min \left\{ \sqrt{1 + \beta_{t-1} \tau_{t-1} L_c} \frac{\|\mathbf{m}^{(t)} - \mathbf{m}^{(t-1)}\|_2}{\|\nabla \log[p(\mathbf{m}^{(t)})] - \nabla \log[p(\mathbf{m}^{(t-1)})]\|_2}, L_c \right\} \quad (28)$$

$$\beta_t = \frac{\tau_t}{\tau_{t-1}} \quad (29)$$

- Step 5. Else reject $\mathbf{m}^{(t+1)} = \mathbf{m}^{(t)}$. Then, repeat this process until the convergence.

The resulting Lip-MALA algorithm follows the same steps as MALA but dynamically adjusts τ , improving stability and sampling efficiency in high dimensions.

3.5 Algorithm Comparison and Tuning Considerations

Table 1 summarizes the key characteristics, advantages, and limitations of the four algorithms, supported by theoretical and empirical studies from the literature.

3.5.1 Practical Implementation Notes

- 300 For our seismic inversion experiments, all algorithms were implemented with the following considerations:
1. Gradient Computation: The gradient $\nabla \log p(\mathbf{m}) = -\nabla S(\mathbf{m})$ was computed analytically, ensuring computational efficiency.
 2. Step Size Tuning: Each algorithm's step size parameter was tuned during a preliminary burn-in phase to achieve target acceptance rates.
 - 305 3. Convergence Diagnostics: We employed diagnostics using Effective Sample Size (ESS) and multivariate ESS (Vats et al., 2019)
 4. Computational Resources: All experiments were conducted on a workstation with Intel Core i9-10900K, 64GB RAM, without GPU acceleration. Execution times were measured as wall-clock time for fair comparison.

The following sections present application of these algorithms to synthetic and real seismic data, evaluating their performance
310 in terms of accuracy, uncertainty quantification, and computational efficiency.

4 Results

4.1 Synthetic test

We test our algorithms using noise-free synthetic seismic traces that were obtained from real data of V_p , V_s and ρ for which synthetic seismic traces were generated from the equation 28 for the angles $\theta_1 = 9^\circ$, $\theta_2 = 18.5^\circ$ and $\theta_3 = 27.5^\circ$ and these
315 synthetic seismic traces will be our observed data. We ran the sampling algorithms described in section 3, producing a large chain of realizations, starting from a prior model configuration corresponding to a low frequency model of V_p , V_s and ρ .

Figure 3 shows the objective function variation curves for the different sampling algorithms. Each iteration involves randomly perturbing the velocities and density of a subset of layers and recalculation of seismic traces. The vertical axis represents the objective function calculated using Equation 5. The horizontal axis shows the number of steps in the Markov chain, each
320 associated with an accepted or rejected perturbation of the velocity and density configuration. The first stage of the chain, associated with the initial configuration and large residues, is called the burn-in or warm-up stage. After subtracting the residuals, the model realizations of velocities and densities satisfactorily explain the seismic data within the data errors. This is called the sampling phase. Realizations produced during the sampling phase are treated as samples from the probability density.

The model settings were modified during the sampling phase, but remain within the probability function, as shown in Figure 4. Figure 5 shows all realizations taken (gray area) in the chain sampling phase for the different algorithms tested in this work, all adjusting the observed seismic data and within the uncertainties of the data. These realizations indicate the features and variability of the velocities and density. Table 2 shows the statistical parameters of mean and standard deviation (SD) which we will compare then with the data obtained from the inference in the different algorithms used.

Our chain sampling phase yielded 10,000 realizations. From these realizations, we calculated the expected values and marginal probabilities of P-wave and S-wave velocities and density as a function of two-way reflection time. These calculations were based on averaging the model performances over the sampling phase. Figure 6 presents the marginal cumulative probability distributions for P and S wave velocities and density, as estimated by the inversion, along with the actual P and S wave velocities and density of the synthetic test. The figure demonstrates the successful prediction of the actual values for all tested algorithms, accurately identifying the main stratification characterized by high and low velocities and the corresponding high and low density.

Table 3 summarizes the performance of the different algorithms tested in predicting P-wave and S-wave velocities and density. The mean, Standard Deviation (SD), correlation, and Root Mean Squared Error (RMSE) are presented for each parameter. The mean and standard deviation values indicate that the predicted values are closely aligned with the true values. Regarding correlation, MH exhibits the lowest correlation for velocity prediction, while HMC achieves the highest. For density prediction, MH and HMC show correlations below 0.29, while MALA and Lip-MALA achieve correlations above 0.60. In terms of RMSE, MH demonstrates the highest error for velocity prediction, while HMC achieves the lowest. For density prediction, MH and HMC exhibit errors above 75.75, while MALA and Lip-MALA maintain errors below 51.04.

Table 4 presents various performance parameters, including acceptance rate and total execution time. Lip-MALA exhibits the highest acceptance rate, while HMC exhibits the lowest. Conversely, MH boasts the lowest total execution time, while HMC demonstrates the highest.

Finally, the convergence of the samples was analyzed a posteriori of the unknown parameters (seismic data parameters) obtained from the different algorithms used. The multivariate effective sample size (mESS) statistic was used. The mESS is a measure that determines the size of an independent and identically distributed sample with the same covariance structure as the sample obtained from an MCMC method for the multivariate case- If we want to know if the chain converges by we can calculate minimum effective sample size (minESS) so that if $mESS > minESS$ we say that the chain converges, if the reader is recommended to review Vats et al. (2019) to delve deeper into the convergence test used in this work. Table 5 shows the summary of mESS and minESS obtained for each method.

4.2 Application to real data

To demonstrate the effectiveness of the algorithms, we applied them to a real dataset of an oil field in eastern Venezuela. The site is located in a formation dominated by clastic rocks, a type of sedimentary rock characterized by alternating layers of sand and shale. The fluids in the pore spaces of these rocks are brine water and oil, without gas. As a preliminary step, we upscaled the P-wave and S-wave velocities obtained from well log data to the corresponding seismic scale using a bandpass filter. This process ensures that the velocity data is consistent with the frequency range of seismic waves. Table 6 presents the descriptive statistics, including mean and standard deviation (SD), for the real data. These values will serve as a baseline for comparison with the results obtained from the inference procedures employed by the various algorithms under consideration.

The seismic traces were obtained from partial stacks for the angles $\theta_1= 19^\circ$, $\theta_2= 24^\circ$, and $\theta_3= 29^\circ$. Utilizing V_p , V_s and ρ logs in seismic scale and wavelets were extracted from the partial stacked seismic data using the frequency content of the data, the synthetic trace was generated using equation (13). The synthetic trace obtained was correlated with observed traces for seismic well tie (see figure 7) obtaining a correlation value of 0.55. The sampling algorithms described in section 3 were implemented, generating a large chain of realizations starting from a prior model configuration corresponding to a low-frequency model of V_p , V_s and ρ .

As depicted in figure 8, the objective function variation curves for each sampling algorithm are presented. During each iteration, a subset of layers undergoes a random perturbation of their velocities and density, followed by a recalculation of the seismic trace. The objective function, calculated using equation 5, is represented on the vertical axis, while the horizontal axis represents the number of steps in the Markov chain. Each step corresponds to an accepted or rejected perturbation of the velocities and density configuration.

The model settings were adjusted during the sampling phase, ensuring they remained within the probability function (Figure 9). Figure 10 illustrates all realizations sampled (gray area) in the chain sampling phase for the various algorithms tested in this study, all of which align with the observed seismic data and fall within the data's uncertainty bounds. These realizations highlight the characteristics and variability of the velocities and density.

Employing a chain sampling scheme, we generated 9,000 realizations from which we extracted the expected values and marginal probabilities of P-wave and S-wave velocities and density, all as functions of two-way reflection time. These calculations were derived by averaging the model performances across the sampling phase. Figure 11 depicts the marginal cumulative probability distributions for P and S wave velocities and density, as inferred from the inversion process, alongside the actual P and S wave velocities and density of the synthetic test.

Table 7 summarizes the performance of the tested algorithms in predicting P-wave and S-wave velocities and density. The mean, Standard Deviation (SD), correlation, and Root Mean Squared Error (RMSE) are presented for each parameter. The predicted values closely align with the true values as evidenced by the mean and standard deviation values. MH exhibits the lowest correlation for velocity prediction, while Lip-MALA achieves the highest. For density prediction, MH and HMC show

390 correlations below 0.28, while MALA and Lip-MALA achieve correlations above 0.48. MH demonstrates the highest error for velocity prediction, while Lip-MALA achieves the lowest. For density prediction, MH and HMC exhibit errors above 151.41, while MALA and Lip-MALA maintain errors below 122.22.

Table 8 presents various performance parameters, including acceptance rate and total execution time. Lip-MALA exhibits the highest acceptance rate, while HMC exhibits the lowest. Conversely, MH boasts the lowest total execution time, while HMC demonstrates the highest.

395 And a final step, as in the synthetic data, was to test the convergence of the chains, this study employed a posteriori analysis to assess the convergence of samples obtained for the unknown seismic data parameters (denoted by m) using various algorithms. The multivariate effective sample size (mESS) statistic served as the convergence metric. The mESS quantifies the equivalent size of an independent and identically distributed (iid) sample possessing the same covariance structure as the sample generated by a Markov Chain Monte Carlo (MCMC) method in the multivariate case.

400 To formally determine chain convergence, a minimum effective sample size (minESS) threshold can be established. If the mESS value surpasses the minESS threshold, convergence is achieved. For a more in-depth exploration of the convergence test employed in this work, readers are referred to Vats et al. (2019). Table 9 summarizes the mESS and minESS values obtained for each method.

4.3 Two-dimensional test with real data

405 In line with the main objective of this study, which is to comparatively evaluate the performance of different MCMC algorithms applied to prestack seismic inversion in geological contexts of varying complexity, this section includes an additional test using real data in a two-dimensional setting. This extension allows us to validate the scalability and robustness of the methods beyond idealized one-dimensional cases, providing evidence of their practical applicability in environments with greater structural and lithological heterogeneity, typical of exploration in real reservoirs.

410 The study area corresponds to a sector of the Eastern Basin of Venezuela, characterized by clastic lithology with alternating sandstones and shales. The seismic data used was acquired using conventional reflection techniques and subsequently reprocessed to generate prestack gathers organized by incidence angle. Three partial stacks were selected, corresponding to central angles of approximately 19° , 24° , and 29° , and used in the inverse process.

The two-dimensional inversion grid was designed with a vertical resolution of 351 samples, corresponding to the two-way time axis from 850 to 1550 ms, and a horizontal extent of 136 cells from 0 to 340 m. Three fundamental elastic parameters were considered: V_p , V_s and ρ . The prior model was built by integrating the structural interpretation of horizons and faults obtained from seismic data, together with low-frequency interpolation of the elastic properties V_p , V_s and ρ . derived from the well logs available in the study area. This approach allowed the establishment of a consistent geological model that served as an initial reference for the Bayesian inversion process.

420 The four MCMC sampling algorithms analyzed were applied in 1D: MH, HMC, MALA, and Lip-MALA. The resulting realizations were used to calculate the posterior marginal distributions for each parameter, as well as their summary statistics (mean and standard deviation).

Quantitative analysis of the figures 12 to 15 reveals important differences between the algorithms. The HMC method presents
425 the lowest standard deviations for all three parameters, indicating greater accuracy and lower uncertainty in the estimation, although at the cost of higher computational costs. MALA and Lip-MALA also offer robust results, with Lip-MALA demonstrating a better compromise between accuracy and computational efficiency, especially in density estimation. In contrast, the MH algorithm shows the greatest uncertainties, confirming its limitations in contexts of higher dimensionality and complexity. From an applied perspective, these results are relevant for choosing the most appropriate method in exploratory contexts. In
430 situations where high accuracy is required and sufficient computational resources are available, HMC may be the preferred option. On the other hand, in scenarios where computational time is a constraint, methods such as Lip-MALA offer an efficient and stable alternative. Density (ρ) remains the most challenging parameter to accurately recover, especially for MH and HMC, which is consistent with the results obtained in the one-dimensional case.

Finally, the incorporation of this two-dimensional test with real data demonstrates the practical applicability of the evaluated
435 algorithms, not only under controlled conditions but also in situations representative of real-world geology. Furthermore, it reinforces the validity of the study's overall conclusions by confirming that gradient-based algorithms are more appropriate for complex inverse scenarios, and that algorithm selection must consider both the level of uncertainty and computational cost.

5 Discussion

This study presents a comparative analysis of Markov Chain Monte Carlo (MCMC) methods for estimating elastic properties
440 from seismic amplitudes. We demonstrate the application of these methods in a field case, employing the following assumptions: (1) a one-dimensional reservoir model represented by stacked seismic traces, (2) seismic data simulation using the small reflectivity approximation, and (3) the Aki-Richards equation for weak contrast to establish the relationship between seismic data and elastic parameters. Notably, the proposed general formulation transcends these assumptions, allowing for the integration of more sophisticated seismic simulation techniques and comprehensive petrophysical models within a similar framework.
445 Why does Lip-MALA yield smaller errors in the theoretical (1D) test while HMC performs better in the two-dimensional real case? In our experiments, two factors explain this: (i) the evaluation metric and (ii) the dimensionality/structure of the posterior.

(i) Metric. In 1D (synthetic and real), we compared primarily RMSE along the trace; Lip-MALA, by enforcing a Lipschitz-based step control, stabilizes local moves and reduces pointwise error (RMSE) in V_p and V_s . In 2D, we mainly
450 reported posterior SD per parameter; HMC, via long Hamiltonian trajectories, explores connected valleys of the

posterior more efficiently and achieves lower global uncertainty (SD), even if RMSE 2D is not emphasized in the main text.

(ii) Dimensionality/heterogeneity. In real 2D, lateral couplings and heterogeneity induce a rough/multimodal posterior; HMC traverses basins more effectively than local Lip-MALA steps, reducing posterior variance.

455 The four methods studied demonstrate acceptable performance, but in-depth analysis reveals notable differences:

- Velocity estimation: In both the synthetic and real-world scenarios, methods that incorporate gradient calculations (HMC, MALA, and Lip-MALA) outperform MH in estimating velocities.
- Density estimation: Density estimation proves to be the most challenging parameter, with MH and HMC exhibiting unsatisfactory results. However, MALA and Lip-MALA showcase more promising performance.
- Execution time: A significant difference emerges in execution time between methods. MH and MALA exhibit shorter execution times compared to HMC and Lip-MALA, which are considerably more time-consuming.

460

465

A natural progression of this research would be to invert prestack seismic data to extract additional elastic parameters and reservoir properties, revealing a more comprehensive subsurface understanding. Similarly, incorporating well log conditioning into the model holds promise, as it could enhance vertical resolution near wells and guarantee that the model aligns with well data at drilling locations.

470

From an applied perspective, the results obtained in this study are relevant for decision-making in real-world exploration settings. For example, in frontier areas with poor well control, the MH algorithm could be used as a rapid evaluation tool due to its low computational cost, albeit with accuracy limitations. In contrast, methods such as HMC or Lip-MALA would be more suitable for mature fields where higher fidelity in estimating elastic properties is required, despite their greater computational demand. The choice of algorithm should be guided not only by statistical metrics but also by the specific requirements of the geophysical project, the geological setting, and the time and resource constraints available.

475

The results obtained in this work show consistency with previous research. For example, Gebraad et al. (2020) highlights the effectiveness of the HMC algorithm in full-waveform elastic inversion problems, particularly due to its ability to efficiently explore the posterior space. However, unlike their FWI-oriented approach, our AVO inversion results indicate that Lip-MALA achieves a better balance between accuracy and computational cost, particularly in density estimation, which is crucial in clastic media with gradual transitions. Similarly, Izzatullah et al. (2021) demonstrated that Langevin dynamics-based methods, such as MALA and its adaptive variants, are more efficient in high-dimensional spaces, which is reflected in our study by better definition of lithological boundaries in inverted images. These parallels confirm that Langevin-derived methods are viable and robust options in real-world seismic scenarios where efficiency and stability are practical priorities.

480

Although the HMC algorithm presents significantly longer runtimes, these may be acceptable within a seismic interpretation workflow that includes validation and multidisciplinary analysis phases. In contexts where the inversion must be performed in near-real time, such as during well drilling (geosteering), methods such as MH or MALA may be more appropriate despite their lower resolution. Therefore, computation time should not be evaluated in isolation, but rather based on operational

485 priorities, available computational resources, and the criticality of the information to be estimated at each stage of the geo-physical project.

6 Conclusions

This study compares various pre-stack inversion methods under an MCMC framework for the estimation of elastic parameters. We invert pre-stacked seismic data to infer velocities (V_P and V_S) and density (ρ), which are linked to the seismic data via the Aki-Richards equation. All methods employed effectively handle the inherent uncertainties associated with seismic and elastic data.

490 The proposed algorithms allow estimating several important aspects of the posterior distribution, such as the means and standard deviations of the posterior parameters. We rigorously validated the algorithms by measuring the quality of the MCMC sample through correlations, plotting the objective function, seismic traces and estimating the RMSE.

The four methods evaluated in this study exhibit acceptable performance overall, but a closer examination reveals notable differences in their specific capabilities. Velocity estimation: In both the simulated and real-world scenarios, methods that leverage gradient calculations (HMC, MALA, and Lip-MALA) demonstrate superior performance in estimating velocities compared to MH. Density estimation: Density estimation poses the most significant challenge, with MH and HMC exhibiting unsatisfactory results. However, MALA and Lip-MALA demonstrate more promising performance in this area. Execution time: A clear distinction emerges in execution time between the methods. MH and MALA exhibit significantly shorter execution times compared to HMC and Lip-MALA, which are considerably more time-consuming.

500 Furthermore, the results of the two-dimensional test with real data showed that in situations where high accuracy is required and sufficient computational resources are available, HMC may be the preferred option. On the other hand, in scenarios where computational time is a constraint, methods such as Lip-MALA offer an efficient and stable alternative. This validation in a context closer to real geology strengthens the study's conclusions. The choice of algorithm must consider not only statistical metrics but also the geophysical context, resource availability, and project purpose. In summary we have:

- Balance runtime vs accuracy: HMC yields lower SD at higher cost; Lip-MALA provides strong local accuracy efficiently.
- Choose metrics deliberately: report both RMSE (fit) and SD (uncertainty) to avoid metric-induced contradictions.
- Use HMC for 2D (and higher) problems to reduce global posterior SD and traverse multi-basin landscapes.
- Use MALA/Lip-MALA to stabilize density estimation and reduce pointwise errors.
- Use Lip-MALA when local accuracy (lower RMSE) in VP/VS is prioritized on 1D settings.

Use of IA tools

515 The authors declare that we use AI tools like ChatGPT and others to improve the writing, structure and make the article more readable for readers.

Competing interests

The authors declare that they have no conflict of interest.

Code and data availability

520 The codes and data for this project are not available due to restrictions imposed by the data provider and the university that funded the project.

Author contribution

RPR and SI designed the study, performed the research, analyzed data, and wrote the paper. GB and RM contributed to refining the ideas, proof the results, carrying out additional analyses, and finalizing this paper.

525 Acknowledgements

The authors would like to thank Yachay Tech University for funding this study through the Research Project MATH24-08. The authors used AI tools to improve the language, readability, and structure of the manuscript. The authors reviewed and edited the output and take full responsibility for the content of the article.

References

530 Aki, K. and Richards, P.: Quantitative Seismology, University Science Books, 2002.

Beskos, A., Pillai, N. S., Roberts, G. O., Sanz-Serna, J.-M., and Stuart, A. M.: Optimal tuning of the hybrid Monte Carlo algorithm, *Bernoulli*, 19, 1501–1534, <https://doi.org/10.3150/12-BEJ414>, 2013.

535 Bosch, M.: The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction, *Geophysics*, 69, 1272–1282, <https://doi.org/10.1190/1.1801944>, 2004.

- Bosch, M., Cara, L., Rodrigues, J., Navarro, A., and Díaz, M.: A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes, *Geophysics*, 72, O29–O39, <https://doi.org/10.1190/1.2783766>, 2007.
- 540
- Buland, A. and Omre, H.: Bayesian linearized AVO inversion, *Geophysics*, 68, 185–198, <https://doi.org/10.1190/1.1543206>, 2003.
- Chib, S. and Greenberg, E.: Understanding the Metropolis-Hastings algorithm, *The American Statistician*, 49, 327–335, <https://doi.org/10.2307/2684568>, 1995.
- 545
- Clochard, V., Delépine, N., Labat, K., and Ricarte, P.: Post-stack versus pre-stack stratigraphic inversion for monitoring purposes: a case study for the saline aquifer of the Sleipner field, *SEG Technical Program Expanded Abstracts 2009*, Society of Exploration Geophysicists, 2417–2421, <https://doi.org/10.1190/1.3255345>, 2009.
- 550
- de Lima, P. D. S., Corso, G., Ferreira, M. S., and de Araújo, J. M.: Acoustic full waveform inversion with Hamiltonian Monte Carlo method, *Physica A: Statistical Mechanics and its Applications*, 617, 128618, <https://doi.org/10.1016/j.physa.2023.128618>, 2023.
- 555
- Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D.: Hybrid Monte Carlo, *Physics Letters B*, 195, 216–222, [https://doi.org/10.1016/0370-2693\(87\)91197-X](https://doi.org/10.1016/0370-2693(87)91197-X), 1987.
- Fichtner, A. and Zunino, A.: Hamiltonian nullspace shuttles, *Geophysical Research Letters*, 46, 644–651, <https://doi.org/10.1029/2018GL080931>, 2019.
- 560
- Gebraad, L., Boehm, C., and Fichtner, A.: Bayesian elastic full-waveform inversion using Hamiltonian Monte Carlo, *Journal of Geophysical Research: Solid Earth*, 125, e2019JB018428, <https://doi.org/10.1029/2019JB018428>, 2020.
- Goodway, B.: AVO and Lamé constants for rock parameterization and fluid detection, *CSEG Recorder*, 26, 39–60, 2001.
- 565
- Hampson, D. P., Russell, B. H., and Bankhead, B.: Simultaneous inversion of prestack seismic data, *SEG Technical Program Expanded Abstracts 2005*, Society of Exploration Geophysicists, 1633–1637, <https://doi.org/10.1190/1.2148008>, 2005.
- Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57, 97–109, <https://doi.org/10.1093/biomet/57.1.97>, 1970.
- 570

- Helland-Hansen, D., Magnus, I., Edvardsen, A., and Hansen, E.: Seismic inversion for reservoir characterization and well planning in the Snorre field, *The Leading Edge*, 16, 269–274, <https://doi.org/10.1190/1.1437616>, 1997.
- 575 Infante, S., Sanchez, L., and Hernández, A.: Stochastic models to estimate population dynamics, *Statistics, Optimization & Information Computing*, 7, 311–328, <https://doi.org/10.19139/soic.v7i2.538>, 2019.
- Izzatullah, M., van Leeuwen, T., and Peter, D.: Accelerated Langevin Dynamics for Bayesian seismic inversion, *Journal of Geophysical Research: Solid Earth*, 125, e2019JB018428, <https://doi.org/10.1029/2019JB018428>, 2020.
- 580 Izzatullah, M., van Leeuwen, T., and Peter, D.: Bayesian seismic inversion: a fast sampling Langevin dynamics Markov chain Monte Carlo method, *Geophysical Journal International*, 227, 1523–1553, <https://doi.org/10.1093/gji/ggab287>, 2021.
- Landa, E. and Treitel, S.: Seismic inversion: What it is, and what it is not, *The Leading Edge*, 35, 277–279, 585 <https://doi.org/10.1190/tle35030277.1>, 2016.
- Landau, L. D. and Lifshitz, E. M.: *Mechanics*, Elsevier Science, 1976.
- Lemons, D. S. and Gythiel, A.: Paul Langevin’s 1908 paper “On the Theory of Brownian Motion” [“Sur la théorie du mouvement brownien,” *C. R. Acad. Sci. (Paris)* 146, 530–533 (1908)], *American Journal of Physics*, 65, 1079–1081, 590 <https://doi.org/10.1119/1.18725>, 1997.
- Ma, X.: Simultaneous inversion of prestack seismic data for rock properties using simulated annealing, *Geophysics*, 67, 1877–1885, <https://doi.org/10.1190/1.1527087>, 2002.
- 595 MacKay, D. J. C.: *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equation of state calculations by fast computing machines, *The Journal of Chemical Physics*, 21, 1087–1092, <https://doi.org/10.1063/1.1699114>, 1953.
- 600 Neal, R. M.: *Bayesian Learning for Neural Networks*, *Lecture Notes in Statistics*, Vol. 118, Springer, <https://doi.org/10.1007/978-1-4612-0745-0>, 1996.

- 605 Neal, R. M.: MCMC using Hamiltonian dynamics, in: Handbook of Markov Chain Monte Carlo, edited by: Brooks, S., Gelman, A., Jones, G. L., and Meng, X.-L., Chapman and Hall/CRC, Boca Raton, 113–162, <https://doi.org/10.1201/b10905-6>, 2012.
- 610 Nemeth, C. and Fearnhead, P.: Stochastic Gradient Markov Chain Monte Carlo, *Journal of the American Statistical Association*, 116, 433–450, <https://doi.org/10.1080/01621459.2020.1847120>, 2021.
- Nemeth, C., Sherlock, C., and Fearnhead, P.: Particle Metropolis-adjusted Langevin algorithms, *Biometrika*, 103, 701–717, 2016.
- 615 Niu, L., Geng, J., Wu, X., Zhao, L., and Zhang, H.: Data-driven method for an improved linearised AVO inversion, *Journal of Geophysics and Engineering*, 18, 1–22, <https://doi.org/10.1093/jge/gxaa065>, 2020.
- Robert, C. P. and Casella, G.: Monte Carlo Statistical Methods, 2nd edn., Springer, New York, <https://doi.org/10.1007/978-1-4757-4145-2>, 2004.
- 620 Roberts, G. O. and Rosenthal, J. S.: Optimal scaling of discrete approximations to Langevin diffusions, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60, 255–268, <https://doi.org/10.1111/1467-9868.00123>, 1998.
- Roberts, G. O. and Tweedie, R. L.: Exponential convergence of Langevin distributions and their discrete approximations, *Bernoulli*, 2, 341–363, <https://doi.org/10.2307/3318418>, 1996.
- 625 Roberts, G. O., Gelman, A., and Gilks, W. R.: Weak convergence and optimal scaling of random walk Metropolis algorithms, *The Annals of Applied Probability*, 7, 110–120, <https://doi.org/10.1214/aop/1034625254>, 1997.
- 630 Roberts, G. and Stramer, O.: Langevin diffusions and Metropolis-Hastings algorithms, *Methodology and Computing in Applied Probability*, 4, 337–357, <https://doi.org/10.1023/A:1023562417138>, 2002.
- Russell, B., Hampson, D., and Bankhead, B.: An Inversion Primer, *CSEG Recorder*, 31, 96–103, 2006.
- 635 Stuart, A. M., Voss, J., and Wilberg, P.: Conditional path sampling of SDEs and the Langevin MCMC method, *Communications in Mathematical Sciences*, 2, 685–697, <https://doi.org/10.4310/CMS.2004.v2.n4.a7>, 2004.

Sánchez, L., Infante, S., Griffin, V., and Rey, D.: Spatio-temporal dynamic model and parallelized ensemble Kalman filter for precipitation data, *Brazilian Journal of Probability and Statistics*, 30, <https://doi.org/10.1214/15-BJPS297>, 2016.

640 Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, <https://doi.org/10.1137/1.9780898717921>, 2005.

Vats, D., Flegal, J. M., and Jones, G. L.: Multivariate output analysis for Markov chain Monte Carlo, *Biometrika*, 106, 321–337, <https://doi.org/10.1093/biomet/asz002>, 2019.

645

Welling, M. and Teh, Y. W.: Bayesian learning via stochastic gradient Langevin dynamics, in: *Proceedings of the 28th International Conference on Machine Learning (ICML 2011)*, Bellevue, WA, USA, 681–688, 2011.

650 Wu, H., Chen, Y., Li, S., and Peng, Z.: Acoustic impedance inversion using Gaussian Metropolis-Hastings sampling with data driving, *Energies*, 12, 2744, <https://doi.org/10.3390/en12142744>, 2019.

655

660

665

670

Table 1: Comparative analysis of MCMC algorithms for seismic inversion

Algorithm	Key Mechanism	Gradient Required	Optimal Acceptance Rate	Tuning Parameters	Computational Cost per Iteration	Best Suited For
MH	Random walk proposal	No	~23.4% (Roberts et al., 1997)	Step size σ	Low ($O(N)$)	Low-dimensional problems, rapid screening
HMC	Hamiltonian dynamics	Yes	~65-80% (Beskos et al., 2013)	Step size ϵ , steps L	High ($O(L \cdot N)$ + gradient)	High-dimensional smooth posteriors
MALA	Langevin diffusion	Yes	~57.4% (Roberts & Rosenthal, 1998)	Step size τ	Moderate ($O(N)$ + gradient)	Moderately high dimensions with gradients
Lip-MALA	Adaptive Langevin	Yes	~50-70% (Izzatullah et al., 2021)	Initial τ , Lipschitz constant	Moderate-High ($O(N)$ + gradient + adaptation)	High-dimensional, heterogeneous posteriors

Table 2: Mean and Standard deviation of elastic parameters used in the synthetic test.

Parameter	Mean	Sd
V_p (m/s)	3068.21	278.29
V_s (m/s)	1553.81	240.60
ρ (Kg/m ³)	2263.57	54.68

675 **Table 3: Statistical parameters for the results obtained for algorithms tested for the synthetic test.**

Parameter	Mean	Sd	Corr	RMSE
MH				
V_p (m/s)	3058.90	394.79	0.64	302.96
V_s (m/s)	1577.99	359.38	0.64	277.26

ρ (Kg/m ³)	2278.00	112.32	0.29	110.39
HMC				
V_p (m/s)	3075.62	312.57	0.90	135.23
V_s (m/s)	1566.98	271.01	0.90	118.15
ρ (Kg/m ³)	2256.42	61.61	0.16	75.75
MALA				
V_p (m/s)	3051.93	326.28	0.85	174.52
V_s (m/s)	1544.13	277.40	0.80	165.30
ρ (Kg/m ³)	2262.27	29.39	0.68	40.64
Lip-MALA				
V_p (m/s)	3062.09	265.42	0.91	112.94
V_s (m/s)	1552.04	217.04	0.89	110.60
ρ (Kg/m ³)	2266.78	59.00	0.60	51.04

Table 4: Other parameters for synthetic test.

Method	Acceptance rate (%)	Total execution time (s)
MH	36.50	24.12
HMC	17.53	9356.99
MALA	25.28	694.54
Lip-MALA	38.49	3337.02

Table 5: Convergence test for synthetic data.

Method	mESS	minESS
MH	8150.89	7458
HMC	8561.10	7458
MALA	7472.03	7458
Lip-MALA	8119.88	7458

680

Table 6: Mean and Standard deviation of elastic parameters used for real data.

Parameter	Mean	Sd
V_p (m/s)	2642.92	249.40
V_s (m/s)	1289.86	205.84
ρ (Kg/m ³)	2180.06	111.89

Table 7: Statistical parameters for the results obtained for algorithms tested for real data.

Parameter	Mean	Sd	Corr	RMSE
MH				
V_p (m/s)	2634.66	255.65	0.64	215.01
V_s (m/s)	1327.43	241.58	0.51	224.74
ρ (Kg/m ³)	2197.22	170.73	0.35	168.44
HMC				
V_p (m/s)	2640.91	199.32	0.69	182.23
V_s (m/s)	1307.19	218.66	0.52	207.86
ρ (Kg/m ³)	2186.65	138.72	0.28	151.41
MALA				
V_p (m/s)	2634.50	217.55	0.65	196.22
V_s (m/s)	1283.90	202.36	0.55	193.84
ρ (Kg/m ³)	2177.61	72.40	0.65	84.42
Lip-MALA				
V_p (m/s)	2642.07	223.19	0.79	155.40
V_s (m/s)	1295.25	175.54	0.75	138.46
ρ (Kg/m ³)	2194.84	125.50	0.48	122.22

685 **Table 8: Other parameters for real data.**

Method	Acceptance rate (%)	Total execution time (s)
MH	32.66	15.48
HMC	3.94	3970.74

MALA	7.38	292.83
Lip-MALA	37.89	1215.87

Table 9: Convergence test for real data.

Method	mESS	minESS
MH	7936.83	7555
HMC	10405.54	7555
MALA	10146.90	7555
Lip-MALA	7979.45	7555

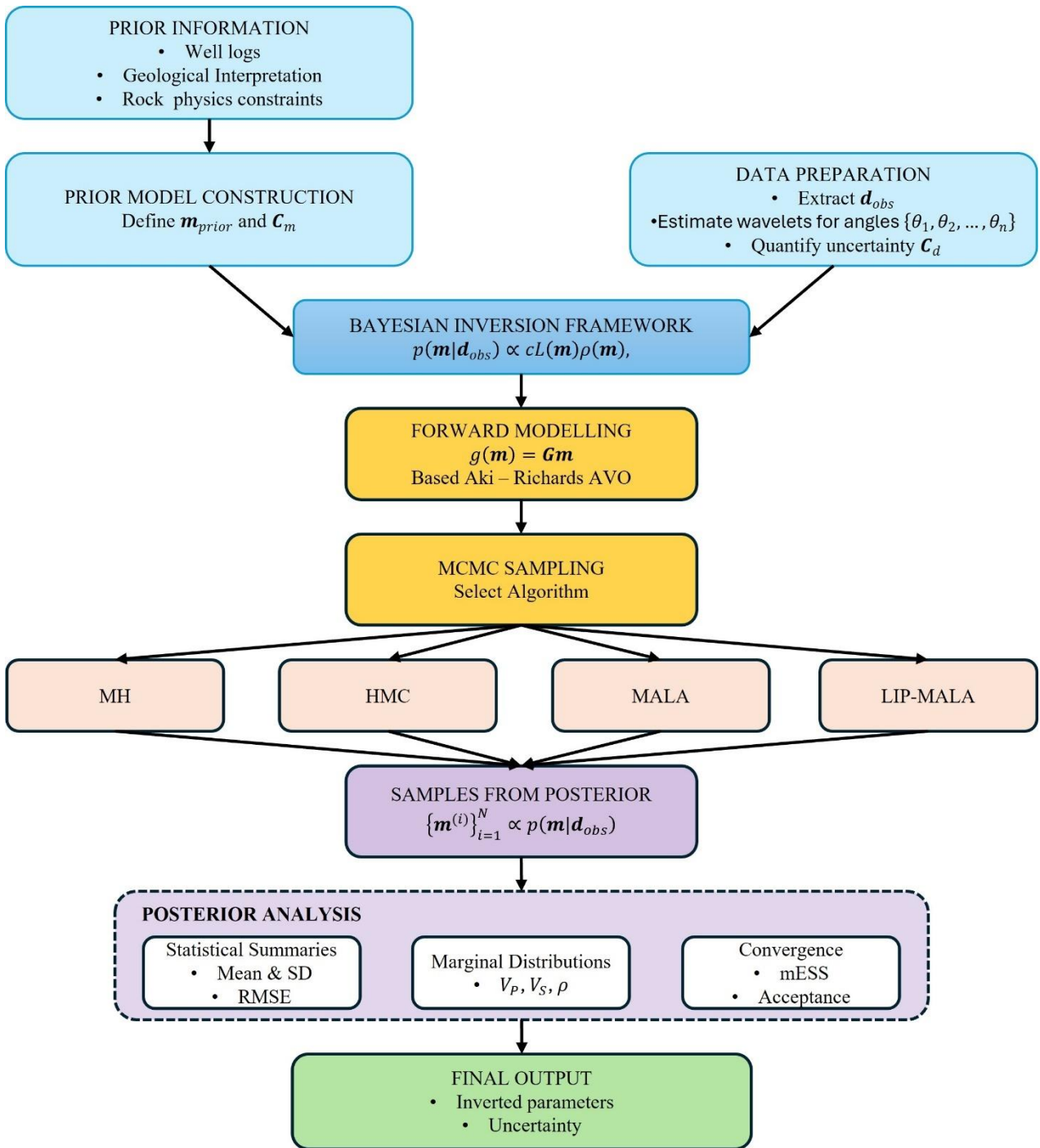
Table 10: Posterior standard deviations (2D real data).

Method	Sd $V_p(m/s)$	Sd $V_s(m/s)$	Sd $\rho(Kg/m^3)$
MH	253.31	234.06	162.53
HMC	170.86	145.72	95.51
MALA	243.09	231.35	109.96
Lip-MALA	195.97	159.14	124.74

690

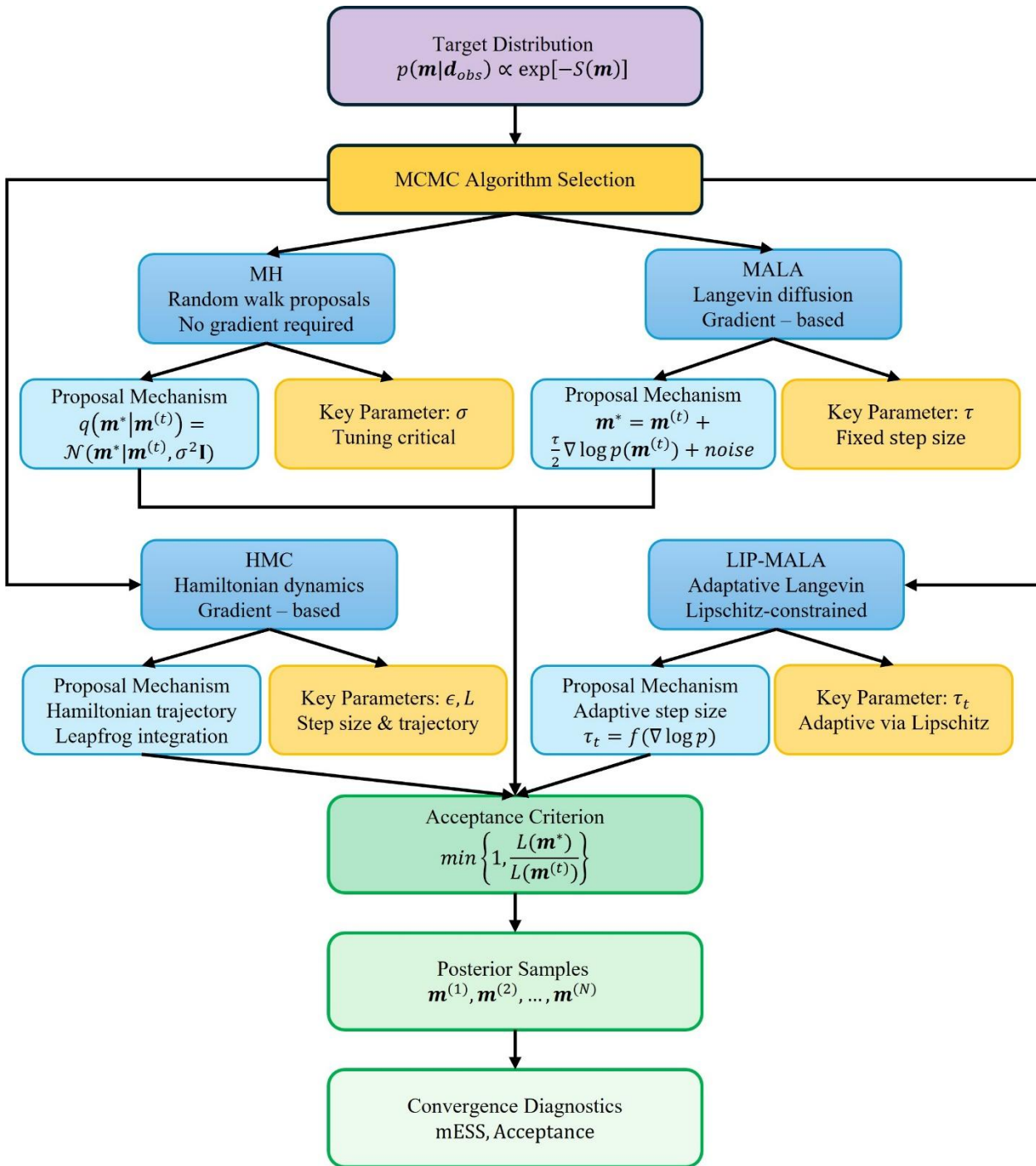
695

700



705

Figure 1: Bayesian seismic inversion workflow implemented in this study, integrating prior information, seismic data, forward modelling, MCMC sampling, and posterior analysis to produce probabilistic estimates of elastic parameters with quantified uncertainties.



710 **Figure 2: Comparative schematic of MCMC algorithms evaluated in this study. All methods generate samples from the posterior distribution but differ in proposal mechanisms, gradient requirements, and parameter tuning strategies.**

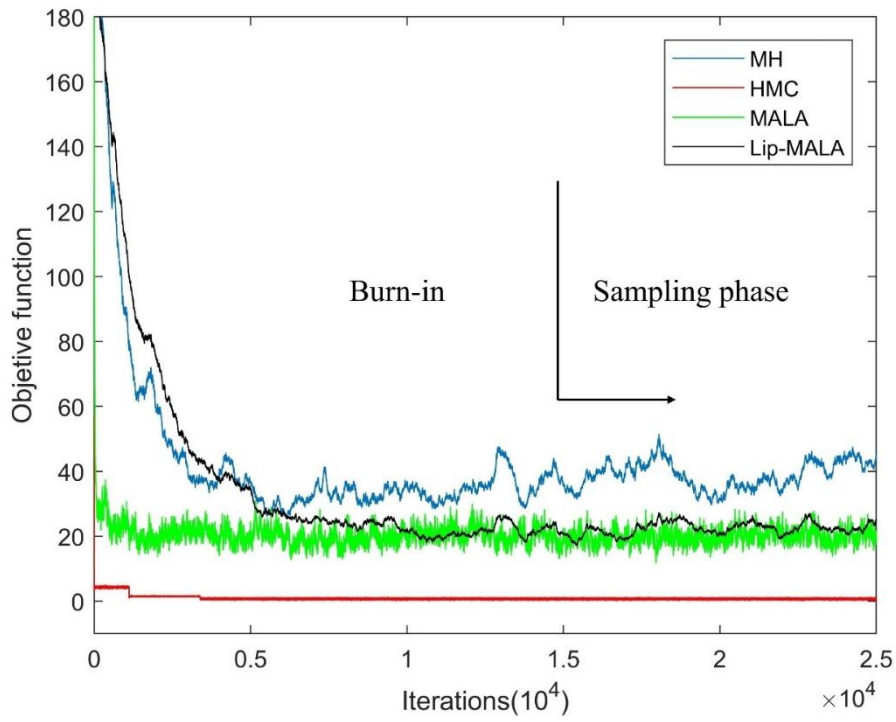
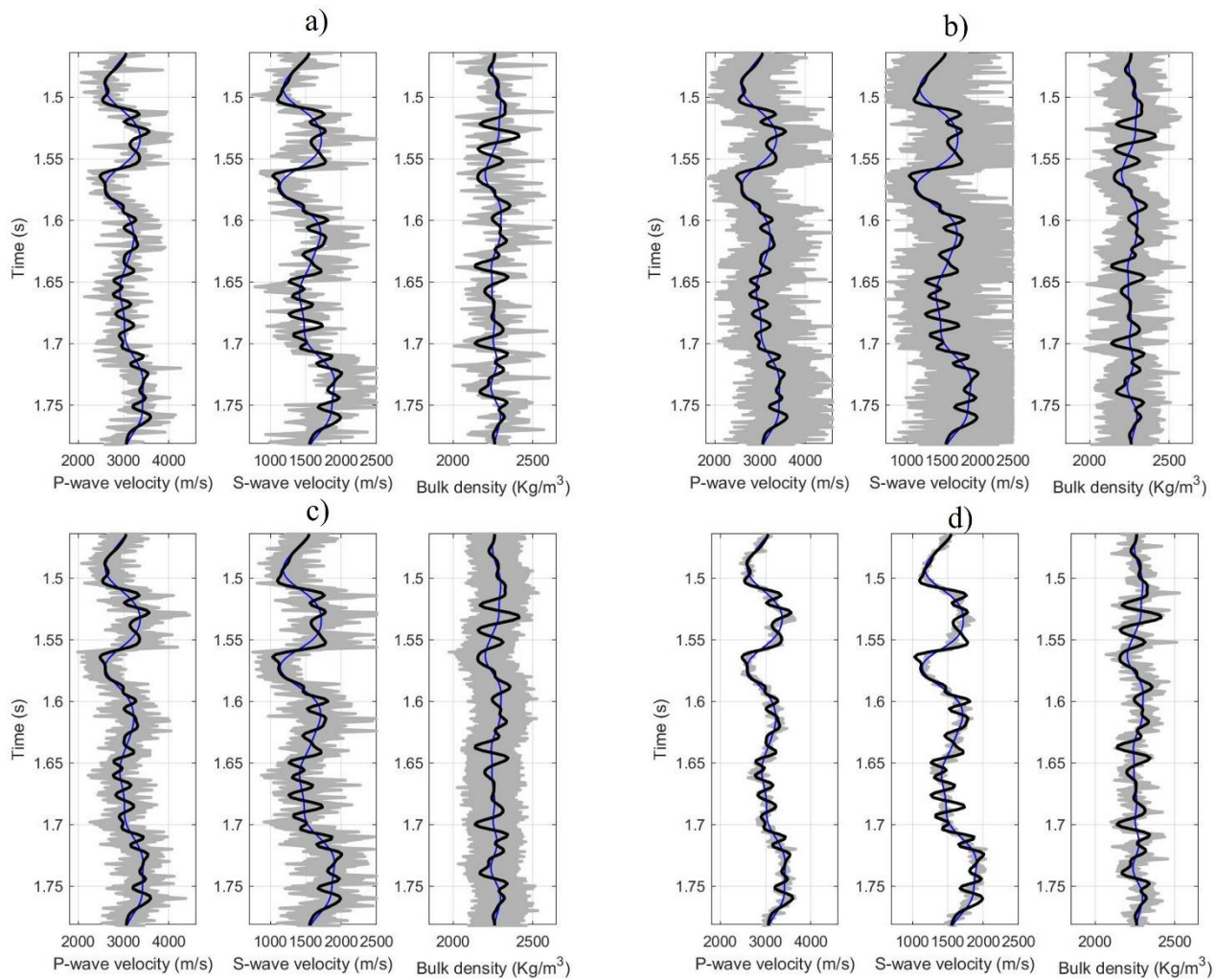


Figure 3: Progress with iterations in the MH (blue line), HMC (red line), MALA (green line) and Lip-MALA (black line) sampling algorithms for synthetic test.



715

Figure 4: True data (black line), prior model (blue line), and accepted realizations of the model (gray) for the synthetic test where a) MH, b) HMC, c) MALA, and d) Lip-MALA.

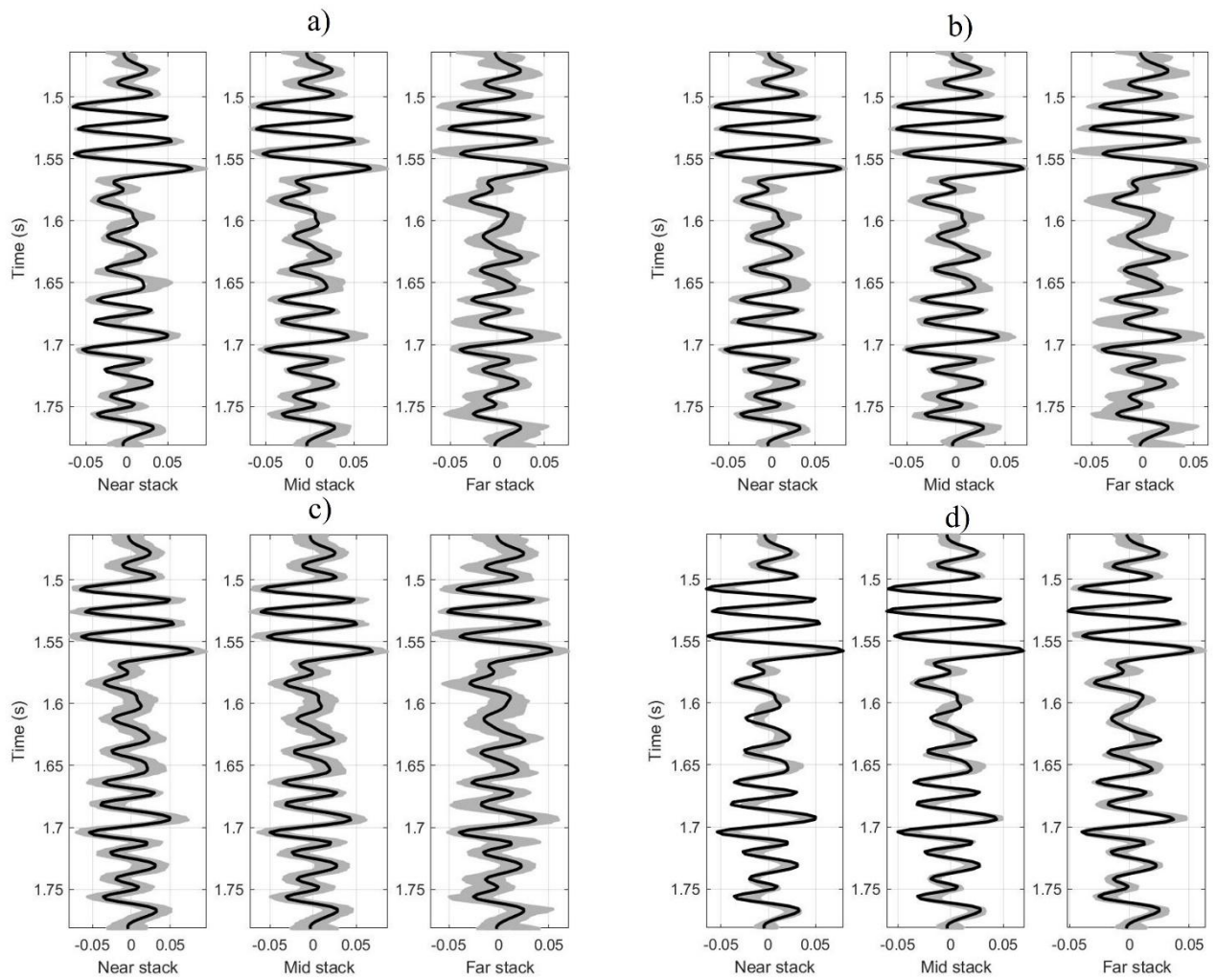


Figure 5: Observed seismic data (black line) and seismic traces obtained from the accepted model realizations (gray) for the synthetic test where a) MH, b) HMC, c) MALA, and d) Lip-MALA.

720

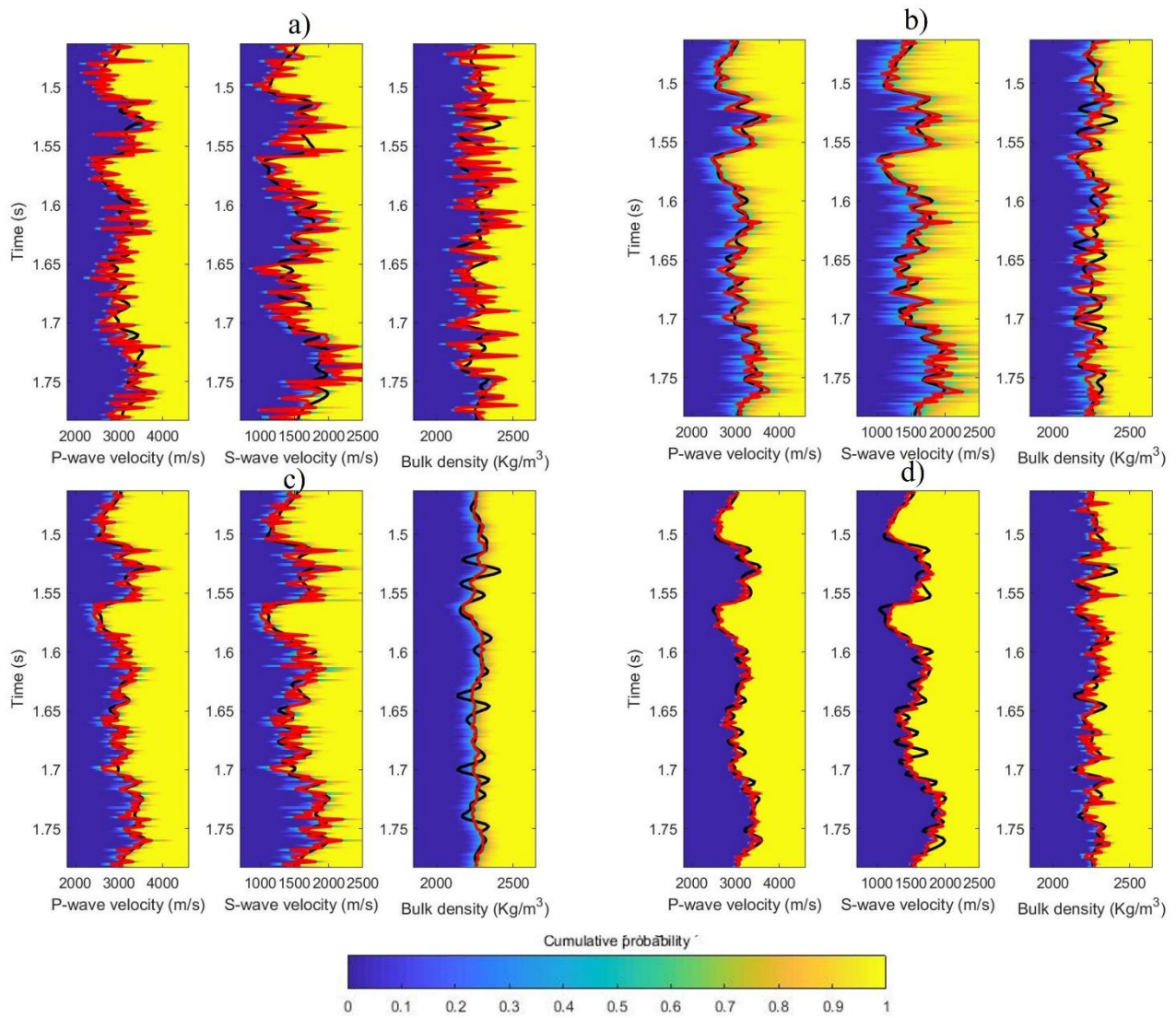
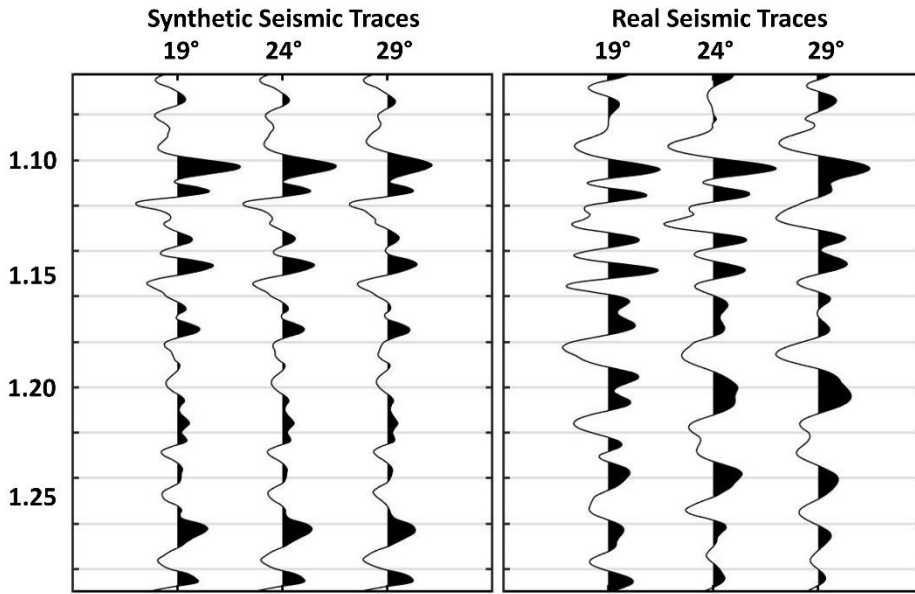


Figure 6: Marginal cumulative probability distributions (color map), true data (black line) and seismic inversion model result (red line) for the synthetic test where a) MH, b) HMC, c) MALA, and d) Lip-MALA.



725 Figure 7: Seismic well tie for real data used.

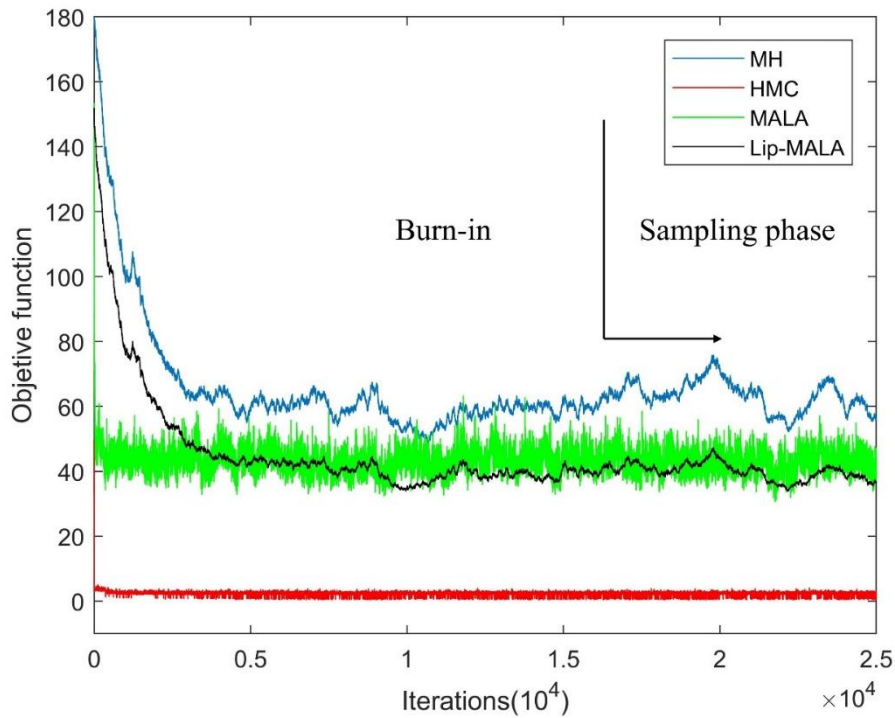
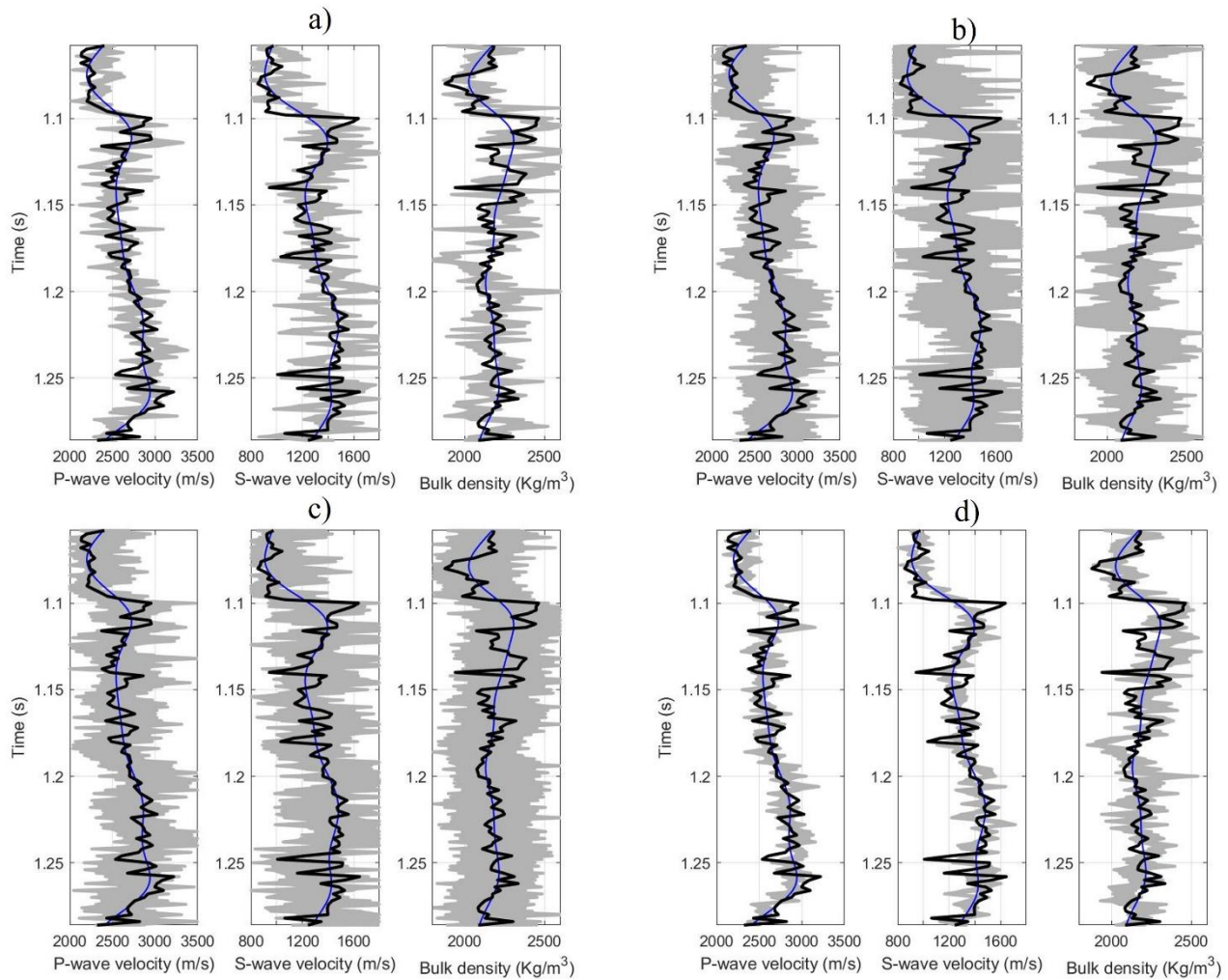


Figure 8: Progress with iterations in the MH (blue line), HMC (red line), MALA (green line) and Lip-MALA (black line) sampling algorithms for real data.



730 **Figure 9: True data (black line), prior model (blue line), and accepted realizations of the model (gray) for real data where a) MH, b) HMC, c) MALA, and d) Lip-MALA.**

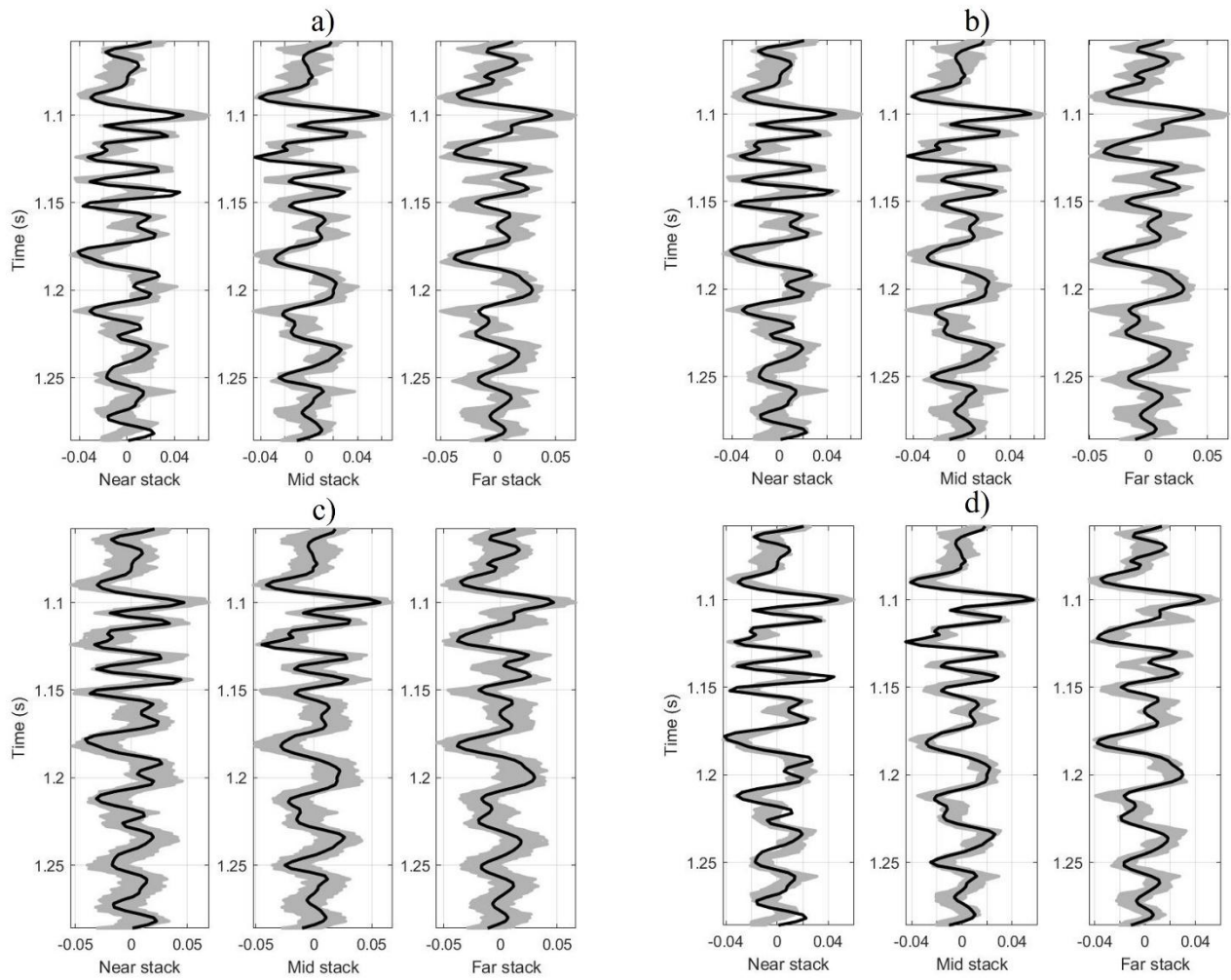
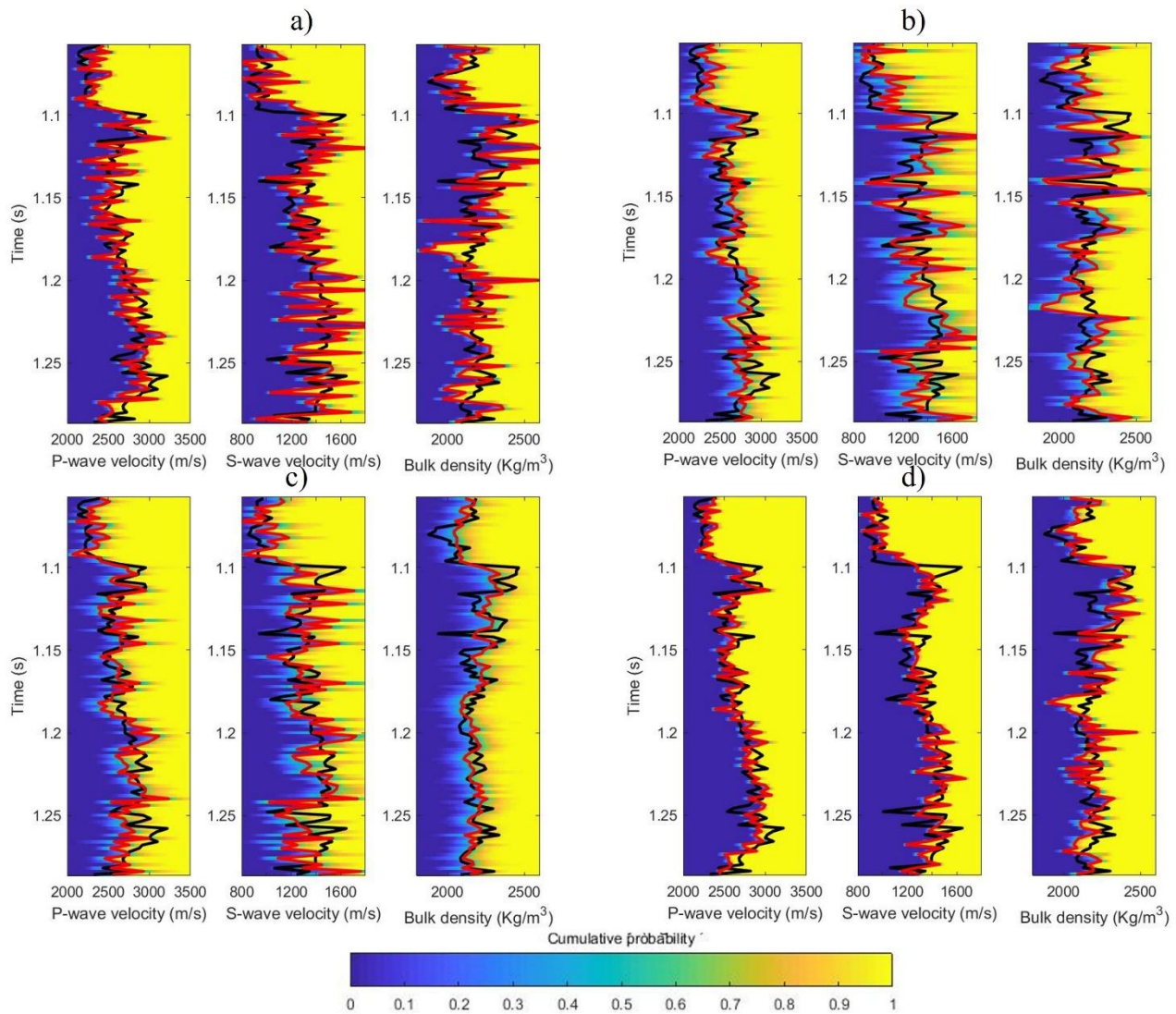


Figure 10: Observed seismic data (black line) and seismic traces obtained from the accepted model realizations (gray) for real data where a) MH, b) HMC, c) MALA, and d) Lip-MALA.



735

Figure 11: Marginal cumulative probability distributions (color map), true data (black line) and seismic inversion model result (red line) for real data where a) MH, b) HMC, c) MALA, and d) Lip-MALA.

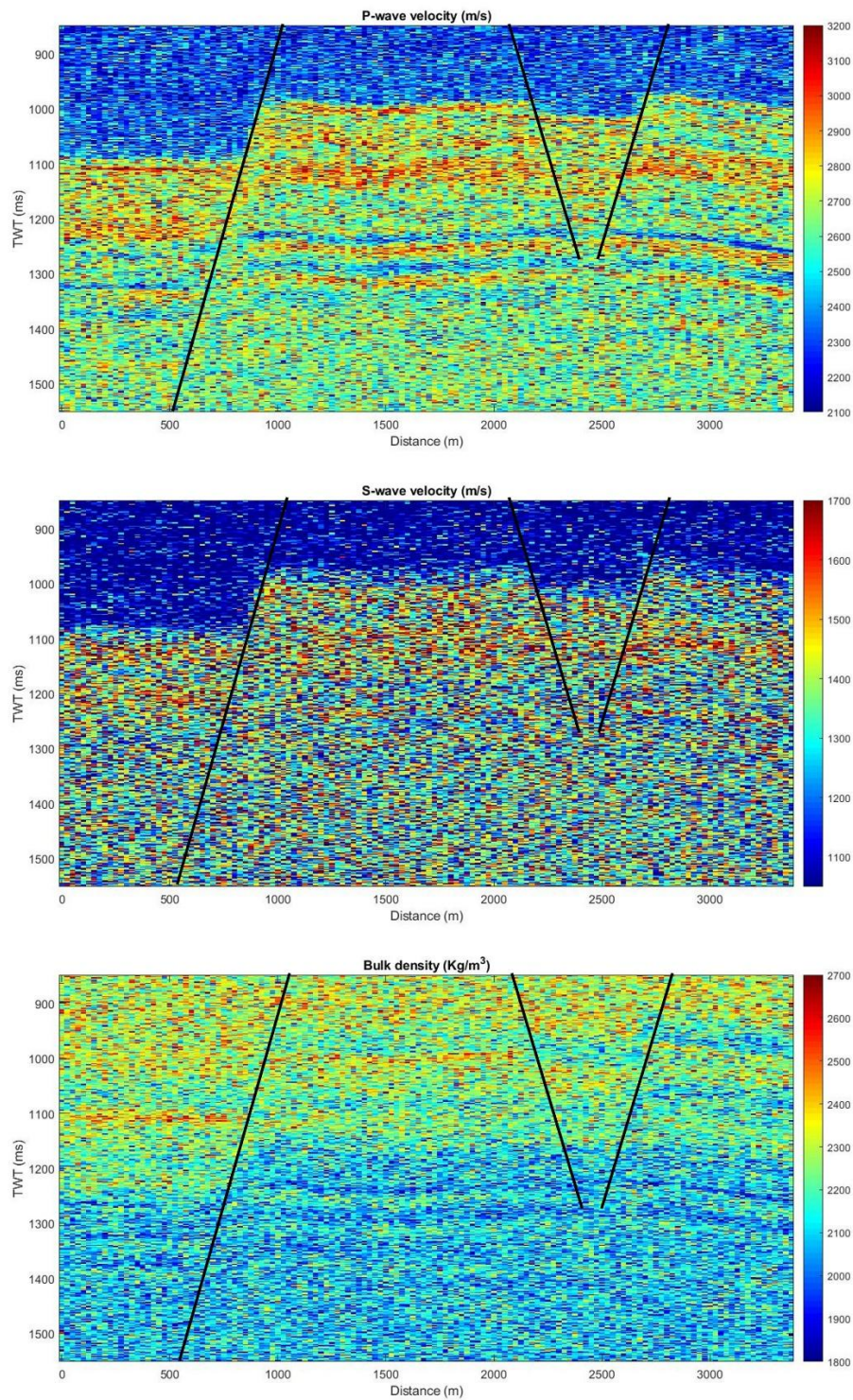
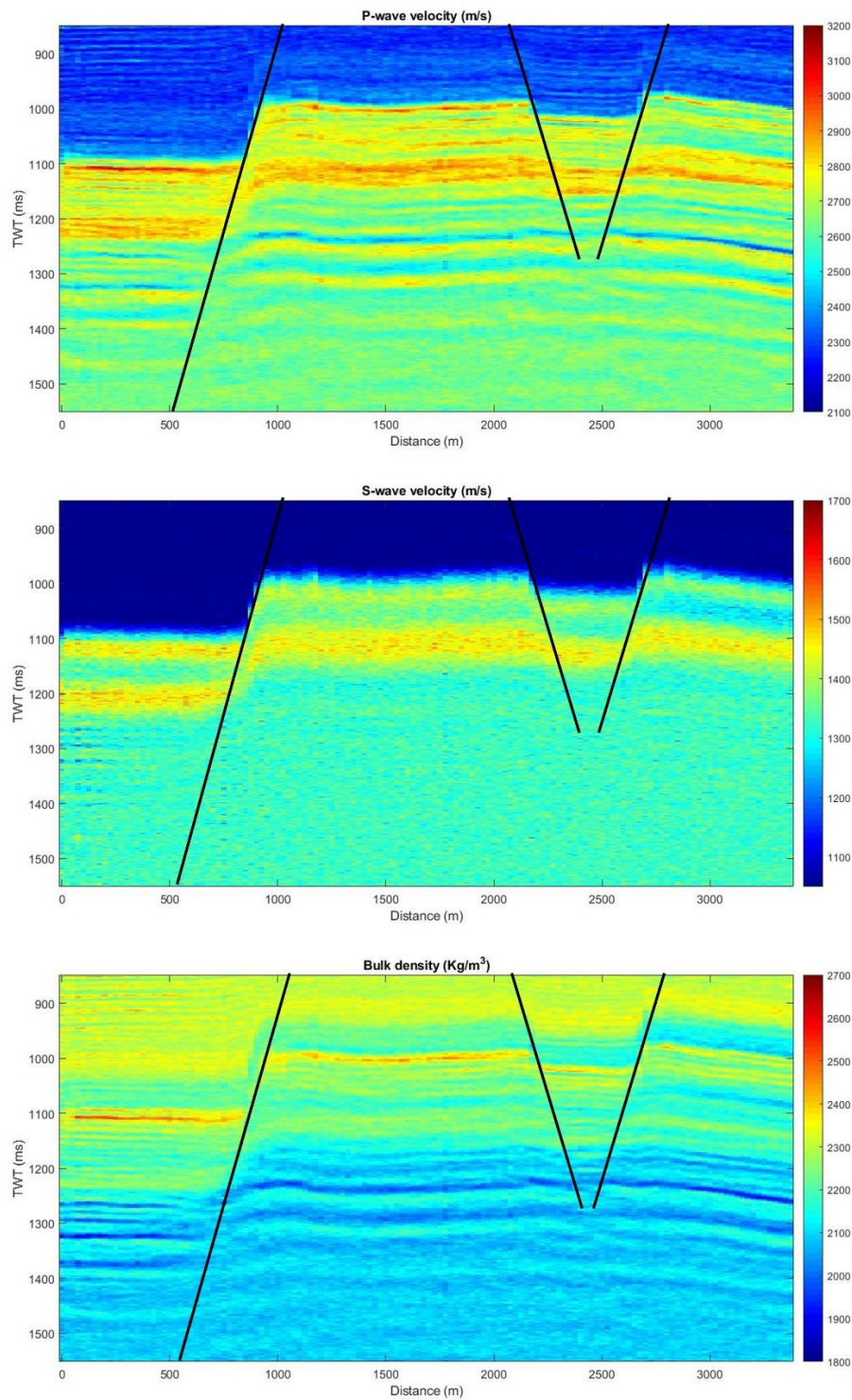


Figure 12: Inverted models of V_p , V_s and ρ obtained using the MH algorithm.



740

Figure 13: Inverted models of V_p , V_s and ρ obtained using the HMC algorithm.

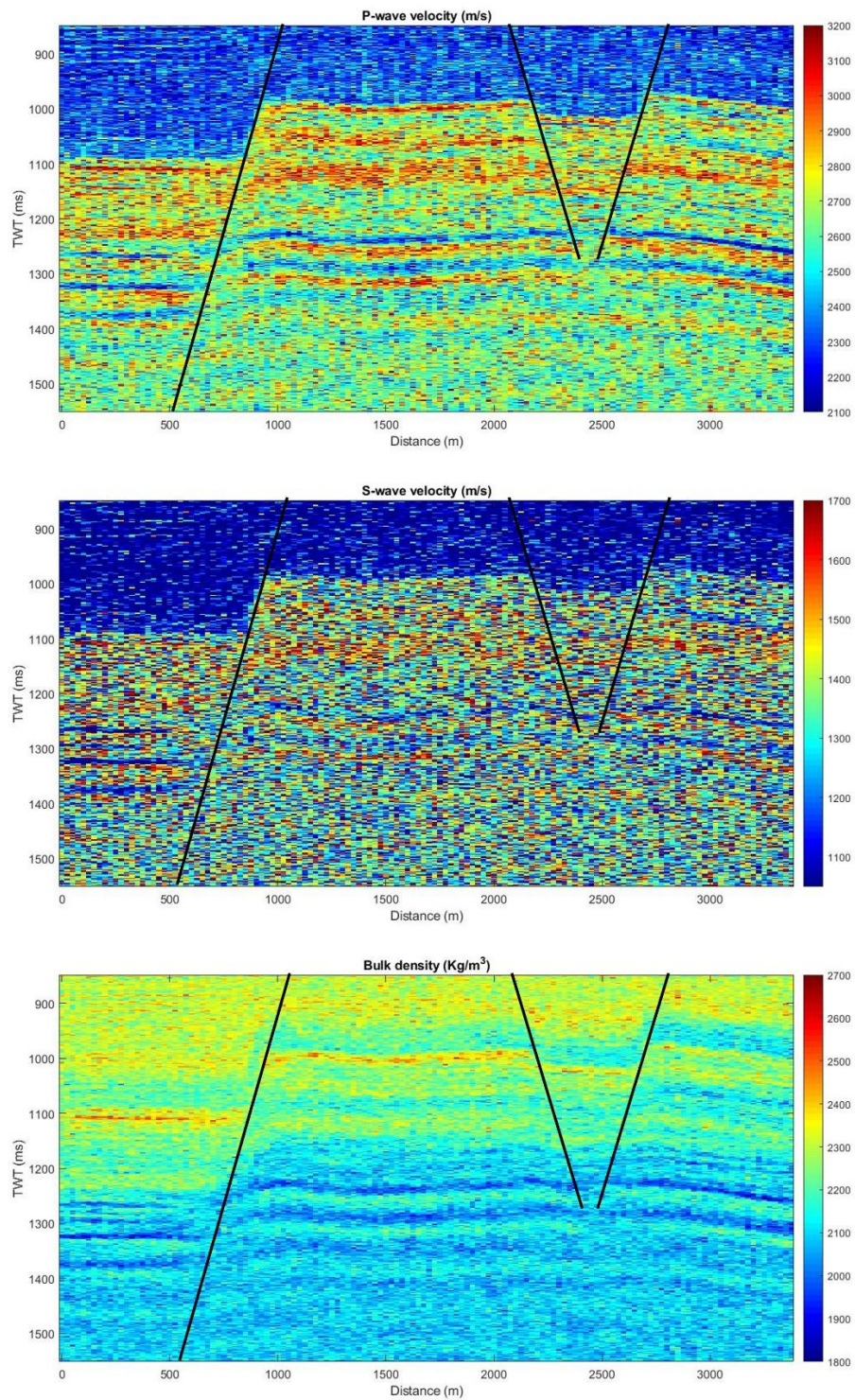
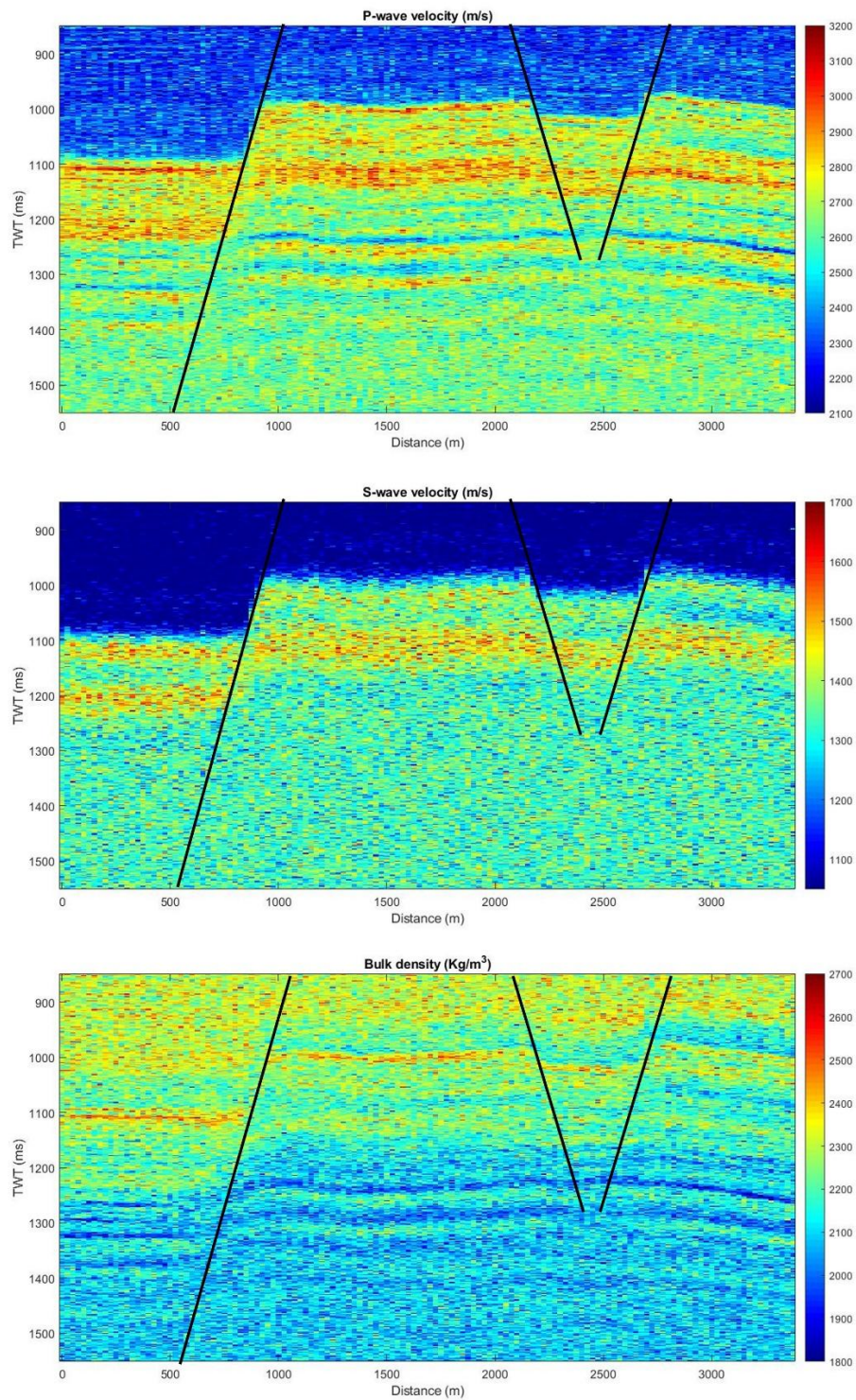


Figure 14: Inverted models of V_p , V_s and ρ obtained using the MALA algorithm.



745 **Figure 15: Inverted models of V_p , V_s and ρ obtained using the Lip-MALA algorithm.**